

Designer antibodies rope immune
cells into attacking cancer p. 930

Ecosystem responses to
invasive plants pp. 934 & 967

Automated synthesis of
whole proteins pp. 941 & 980

Science

\$15
29 MAY 2020
sciencemag.org

AAAS

BACTERIA IN TUMORS

Microbes reside in many
human cancers

pp. 938 & 973

CONTENTS

29 MAY 2020 • VOLUME 368 • ISSUE 6494



NEWS

IN BRIEF

920 News at a glance

IN DEPTH

922 Scientists put survivors' blood plasma to the test

Randomized trials are underway after small studies suggest transfusions can save lives *By K. Kupferschmidt*

923 Doctors race to understand inflammatory condition in kids

Studies are launching to pinpoint children most at risk *By J. Couzin-Frankel*

PODCAST

924 New tools aim to tame pandemic paper tsunami

Researchers harness AI and old-fashioned expertise to glean findings, but face challenges *By J. Brainard*

926 UV radiation blamed in ancient mass extinction

Malformed spores suggest powerful storms drove ozone loss as climate warmed *By P. Voosen*

SCIENCE ADVANCES RESEARCH ARTICLE BY J. E. A. MARSHALL ET AL. 10.1126/SCIADV.ABA0768

927 Core progress in AI has stalled in some fields

When tuned up, old algorithms can match the abilities of their successors *By M. Hutson*

928 Sick chinchillas languish at farms that supply researchers

Years of animal welfare violations documented by USDA *By M. Wadman*

FEATURES

930 Forced into battle

Bispecific antibodies unleash T cells against cancer by physically tethering them to tumor cells *By J. Kaiser*

INSIGHTS

PERSPECTIVES

934 Exotic plants get a little help from their friends

Interactions with herbivores and microbes link exotic-plant success with carbon cycling *By C. Urcelay and A. T. Austin*

RESEARCH ARTICLE p. 967

936 Remote activation of cellular signaling

Electric fields stimulate genetically modified pancreatic β cells to secrete insulin in mice *By M. I. Brier and J. S. Dordick*

RESEARCH ARTICLE p. 993

938 Probing the tumor micro(b)environment

Bacteria are widespread in tumors, are found within cells, and differ by cancer type *By C. E. Atreya and P. J. Turnbaugh*

RESEARCH ARTICLE p. 973

939 Tracking the rapid pace of a retreating ice sheet

Seafloor mapping shows that Antarctic ice sheets retreated faster during the last deglaciation than today *By M. Jakobsson*

REPORT p. 1020

941 Catching up to nature's ribosomes

Flow chemistry enables total chemical synthesis of proteins without solution-phase ligation *By C. Proulx*

RESEARCH ARTICLE p. 980

942 The search for a COVID-19 animal model

A comparison of SARS CoV-2 replication, transmission, and disease in mice to monkeys *By S. S. Lakdawala and V. D. Menachery*

REPORTS pp. 1012 & 1016

943 Killer cells add fire to fuel immunotherapy

Cytotoxic lymphocytes induce inflammatory target cell death, amplifying antitumor immunity *By C. J. Nicolai and D. H. Raulet*

REPORT p. 965

945 Rapid COVID-19 vaccine development

Finding the fastest pathway to vaccine availability includes the avoidance of safety pitfalls *By B. S. Graham*

947 Eleanor Margaret Burbidge (1919–2020)

Astronomer, astrophysicist, and champion of gender equality *By A. Boksenberg*

POLICY FORUM

948 A strategic approach to COVID-19 vaccine R&D

A public-private partnership and platform for harmonized clinical trials aims to accelerate licensure and distribution *By L. Corey et al.*

951 Ethics and governance for digital disease surveillance

The question is not whether to use new data sources but how *By M. M. Mello and C. J. Wang*

BOOKS ET AL.

955 Revisiting rights in an ever-evolving world

Data and genetic policy join proposed privileges for nonhumans, robots, and nature *By D. Greenbaum*

956 The will to act when the data are dire

Then and now, American politicians have failed to heed scientific warnings *By E. L. Milam*

LETTERS

957 Ease restrictions on U.S. blood donations

By C. Vasavda et al.

957 The precarious position of postdocs during COVID-19

By M. A. Ahmed et al.

958 Sumatran rhinoceros on the brink of extinction

By S. S. Lam et al.

958 Technical Comment abstracts

RESEARCH

IN BRIEF

961 From *Science* and other journals

REVIEW

964 Forest ecology

Pervasive shifts in forest dynamics in a changing world *N. G. McDowell et al.*

REVIEW SUMMARY; FOR FULL TEXT:
DX.DOI.ORG/10.1126/SCIENCE.AAZ9463

RESEARCH ARTICLES

965 Immunology

Granzyme A from cytotoxic lymphocytes cleaves GSDMB to trigger pyroptosis in target cells *Z. Zhou et al.*

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT:
DX.DOI.ORG/10.1126/SCIENCE.AAZ7548
PERSPECTIVE p. 943

966 Magnetism

Self-induced spin glass state in elemental and crystalline neodymium *U. Kamber et al.*

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT:
DX.DOI.ORG/10.1126/SCIENCE.AAY6757

967 Invasive species

Biotic interactions drive ecosystem responses to exotic plant invaders *L. P. Waller et al.*

PERSPECTIVE p. 934

973 Cancer

The human tumor microbiome is composed of tumor type-specific intracellular bacteria *D. Nejman et al.*

PERSPECTIVE p. 938

980 Peptide synthesis

Synthesis of proteins by automated flow chemistry *N. Hartrampf et al.*

PERSPECTIVE p. 941

988 Ocean circulation

Strengthening of the Kuroshio current by intensifying tropical cyclones *Y. Zhang et al.*

964

993 Synthetic biology

Electrogenetic cellular insulin release for real-time glycemic control in type 1 diabetic mice *K. Krawczyk et al.*

PERSPECTIVE p. 936

REPORTS

1002 Separations

Control of zeolite pore interior for chemoselective alkyne/olefin separations *Y. Chai et al.*

1007 Organic chemistry

Total synthesis of bryostatin 3 *B. M. Trost et al.*

Coronavirus

1012 Comparative pathogenesis of COVID-19, MERS, and SARS in a nonhuman primate model *B. Rockx et al.*

1016 Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2 *J. Shi et al.*

PERSPECTIVE p. 942

1020 Ice sheets

Delicate seafloor landforms reveal past Antarctic grounding-line retreat of kilometers per year *J. A. Dowdeswell et al.*

PERSPECTIVE p. 939; PODCAST

DEPARTMENTS

919 Editorial

COVID-19 research in Africa

By Linda-Gail Bekker and Valerie Mizrahi

1026 Working Life

You never know until you try *By Tess Torregrosa*

ON THE COVER

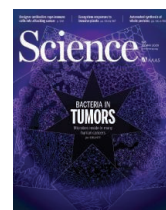


Illustration of bacteria in tissue sections from multiple human tumor types. Bacteria are frequently found in tumors, and different tumor types have distinct microbial compositions, as represented here by the different colors of

bacteria. The image also reflects the observed intracellular localization of tumor bacteria, mostly inside cancer and immune cells. Future studies will explore the role of these bacteria in tumor development and treatment response. See pages 938 and 973. *Illustration: V. Altounian/Science*

Science Staff	916
AAAS News & Notes	959
Science Careers	1025

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2020 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership, including subscription (12 months): \$165 (\$74 allocated to subscription). Domestic institutional subscription (51 issues): \$2148; Foreign postage extra: Air assist delivery: \$98. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request. GST #125488122. Publications Mail Agreement Number 1069624. Printed in the U.S.A. Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. Postmaster: Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090-6178. Single-copy sales: \$15 each plus shipping and handling available from backissues.sciencemag.org; bulk rate on request. Authorization to reproduce material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for Science is 0036-8075. Science is indexed in the Reader's Guide to Periodical Literature and in several specialized indexes.

Editor-in-Chief Holden Thorp, hthorp@aaas.org

Executive Editor Monica M. Bradford

Editors, Research Valda Vinson, Jake S. Yeston Editor, Insights Lisa D. Chong

DEPUTY EDITORS Julia Fahrenkamp-Uppenbrink (UK), Stella M. Hurlley (UK), Phillip D. Szuromi, Sacha Vignieri **SR. EDITORIAL FELLOW** Andrew M. Sugden (UK) **SR. EDITORS** Gemma Alderton (UK), Caroline Ash (UK), Brent Grocholski, Pamela J. Hines, Paula A. Kiberstis, Marc S. Lavine (Canada), Steve Mao, Ian S. Osborne (UK), Beverly A. Purnell, L. Bryan Ray, H. Jesse Smith, Keith T. Smith (UK), Jelena Stajic, Peter Stern (UK), Valerie B. Thompson, Brad Wible, Laura M. Zahn **ASSOCIATE EDITORS** Michael A. Funk, Priscilla N. Kelly, Tage S. Rai, Seth Thomas Scanlon (UK), Yury V. Suleymanov **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Harry Jach, Lauren Kmeck **CONTENT PRODUCTION EDITORS** Amelia Beyna, Jeffrey E. Cook, Chris Filatreau, Julia Katris, Nida Masiulis, Suzanne M. White **SR. EDITORIAL COORDINATORS** Carolyn Kyle, Beverly Shields **EDITORIAL COORDINATORS** Aneera Dobbins, Joi S. Granger, Jeffrey Hearn, Lisa Johnson, Maryrose Madrid, Ope Martins, Shannon McMahon, Jerry Richardson, Hilary Stewart (UK), Alana Warnke, Alice Whaley (UK), Anita Wynn **PUBLICATIONS ASSISTANTS** Jeremy Dow, Alexander Kief, Ronnel Navas, Brian White **EXECUTIVE ASSISTANT** Jessica Slater **ASI DIRECTOR, OPERATIONS** Janet Clements (UK) **ASI SR. OFFICE ADMINISTRATOR** Jessica Waldoock (UK)

News Editor Tim Appenzeller

NEWS MANAGING EDITOR John Travis **INTERNATIONAL EDITOR** Martin Enserink **DEPUTY NEWS EDITORS** Elizabeth Culotta, Lila Guterman, David Grimm, Eric Hand (Europe), David Malakoff **SR. CORRESPONDENTS** Daniel Clerly (UK), Jon Cohen, Jeffrey Mervis, Elizabeth Pennisi **ASSOCIATE EDITORS** Jeffrey Brinard, Catherine Maticic **NEWS REPORTERS** Adrian Cho, Jennifer Couzin-Frankel, Jocelyn Kaiser, Kelly Servick, Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman **INTERNS** Eva Frederick, Rodrigo Perez Ortega **CONTRIBUTING CORRESPONDENTS** Warren Cornwall, Ann Gibbons, Mara Hvistendahl, Sam Kean, Eli Kintisch, Kai Kupferschmidt (Berlin), Andrew Lawler, Mitch Leslie, Eliot Marshall, Virginia Morell, Dennis Normile (Shanghai), Elisabeth Pain (Careers), Charles Pillar, Michael Price, Tania Rabesandratana (Barcelona), Emily Underwood, Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **CAREERS** Donisha Adams, Rachel Bernstein (Editor), Katie Langin (Acting Editor) **COPY EDITORS** Julia Cole (Senior Copy Editor), Cyra Master (Copy Chief) **ADMINISTRATIVE SUPPORT** Meagan Welland

Creative Director Beth Rakouskas

DESIGN MANAGING EDITOR Marcy Atarod **GRAPHICS MANAGING EDITOR** Alberto Cuadra **PHOTOGRAPHY MANAGING EDITOR** William Douthitt **WEB CONTENT STRATEGY MANAGER** Kara Estelle-Powers **SENIOR DESIGNER** Chrystal Smith **DESIGNER** Christina Aycock **GRAPHICS EDITOR** Nirja Desai **INTERACTIVE GRAPHICS EDITOR** Xing Liu **SENIOR SCIENTIFIC ILLUSTRATORS** Valerie Altounian, Chris Bickel **SCIENTIFIC ILLUSTRATOR** Alice Kitterman **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **SENIOR PHOTO EDITOR** Emily Petersen **PHOTO EDITOR** Kaitlyn Dolan

Chief Executive Officer and Executive Publisher Sudip Parikh

Publisher, Science Family of Journals Bill Moran

DIRECTOR, BUSINESS SYSTEMS AND FINANCIAL ANALYSIS Randy Yi **DIRECTOR, BUSINESS OPERATIONS & ANALYSIS** Eric Knott **DIRECTOR OF ANALYTICS** Enrique Gonzales **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **SENIOR BUSINESS ANALYST** Cory Lipman, Meron Kebede **FINANCIAL ANALYST** Alexander Lee **ADVERTISING SYSTEM ADMINISTRATOR** Tina Burks **SENIOR SALES COORDINATOR** Shirley Young **DIGITAL/PRINT STRATEGY MANAGER** Jason Hillman **QUALITY TECHNICAL MANAGER** Marcus Spiegler **ASSISTANT MANAGER DIGITAL/PRINT** Rebecca Doshi **SENIOR CONTENT SPECIALISTS** Steve Forrester, Jacob Hedrick, Antoinette Hodal, Lori Murphy **DIGITAL PRODUCTION MANAGER** Lisa Stanford **CONTENT SPECIALIST** Kimberley Oster **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **DESIGNER, CUSTOM PUBLISHING** Jeremy Huntsinger **SR. TRAFFIC ASSOCIATE** Christine Hall **SPECIAL PROJECTS ASSOCIATE** Sarah Dhere **ASSOCIATE DIRECTOR, BUSINESS DEVELOPMENT** Justin Sawyers **GLOBAL MARKETING MANAGER** Allison Pritchard **DIGITAL MARKETING MANAGER** Aimee Aponte **JOURNALS MARKETING MANAGER** Shawana Arnold **MARKETING ASSOCIATES** Tori Velasquez, Mike Romano, Ashley Hylton **DIGITAL MARKETING SPECIALIST** Asleigh Rojanavongse **SENIOR DESIGNER** Kim Huynh

DIRECTOR AND SENIOR EDITOR, CUSTOM PUBLISHING Sean Sanders **ASSISTANT EDITOR, CUSTOM PUBLISHING** Jackie Oberst

DIRECTOR, PRODUCT & PUBLISHING DEVELOPMENT Chris Reid **DIRECTOR, BUSINESS STRATEGY AND PORTFOLIO MANAGEMENT** Sarah Whalen **ASSOCIATE DIRECTOR, PRODUCT MANAGEMENT** Kris Bishop **SR. PRODUCT ASSOCIATE** Robert Koepke **DIGITAL PRODUCT STRATEGIST** Michael Hardesty **SPJ ASSOCIATE** Samantha Bruno Fuller

DIRECTOR, INSTITUTIONAL LICENSING Iquo Edim **ASSOCIATE DIRECTOR, RESEARCH & DEVELOPMENT** Elisabeth Leonard **MARKETING MANAGER** Kess Knight **SENIOR INSTITUTIONAL LICENSING MANAGER** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGER** Marco Castellani **MANAGER, AGENT RELATIONS & CUSTOMER SUCCESS** Judy Lillibridge **SENIOR OPERATIONS ANALYST** Lana Guz **FULFILLMENT COORDINATOR** Melody Stringer **SALES COORDINATOR** Josh Haverlock

DIRECTOR, GLOBAL SALES Tracy Holmes **EAST COAST AND MID WEST SALES** Stephanie O'Connor **US WEST COAST SALES** Lynne Stickrod **US SALES MANAGER, SCIENCE CAREERS** Claudia Paulsen-Young **US SALES REP, SCIENCE CAREERS** Tracy Anderson **ASSOCIATE DIRECTOR, ROW** Roger Gonçalves **SALES REP, ROW** Sarah Lelarge **SALES ADMIN ASSISTANT, ROW** Bryony Cousins **DIRECTOR OF GLOBAL COLLABORATION AND ACADEMIC PUBLISHING RELATIONS** Asia Xiaoying Chu **ASSOCIATE DIRECTOR, INTERNATIONAL COLLABORATION** Grace Yao **SALES MANAGER** Danny Zhao **MARKETING MANAGER** Kilo Lan **ASCA CORPORATION, JAPAN** Kaoru Sasaki (Tokyo), Miyuki Tani (Osaka) **COLLABORATION/CUSTOM PUBLICATIONS/JAPAN** Adarsh Sandhu

DIRECTOR, COPYRIGHT, LICENSING AND SPECIAL PROJECTS Emilie David **RIGHTS AND LICENSING COORDINATOR** Jessica Adams **RIGHTS AND PERMISSIONS ASSOCIATE** Elizabeth Sandler **CONTRACTS AND LICENSING ASSOCIATE** Lili Catlett

MAIN HEADQUARTERS

Science/AAAS
1200 New York Ave. NW
Washington, DC 20005

SCIENCE INTERNATIONAL

Clarendon House
Clarendon Road
Cambridge, CB2 8FH, UK

SCIENCE CHINA

Room 1004, Culture Square
No. 59 Zhongguancun St.
Haidian District, Beijing, 100872

SCIENCE JAPAN

ASCA Corporation
Sibaura TY Bldg. 4F, 1-14-5
Shibaura Minato-ku
Tokyo, 108-0073 Japan

EDITORIAL

science_editors@aaas.org

NEWS

science_news@aaas.org

INFORMATION FOR AUTHORS

sciencemag.org/authors/
science-information-authors

REPRINTS AND PERMISSIONS

sciencemag.org/help/
reprints-and-permissions

MEDIA CONTACTS

scipak@aaas.org

MULTIMEDIA CONTACTS

SciencePodcast@aaas.org
ScienceVideo@aaas.org

INSTITUTIONAL SALES

AND SITE LICENSES

sciencemag.org/librarian

PRODUCT ADVERTISING

& CUSTOM PUBLISHING
advertising.sciencemag.org/
products-services
science_advertising@aaas.org

CLASSIFIED ADVERTISING

advertising.sciencemag.org/
science-careers
advertise@sciencecareers.org

JOB POSTING CUSTOMER SERVICE

employers.sciencecareers.org
support@sciencecareers.org

MEMBERSHIP AND INDIVIDUAL

SUBSCRIPTIONS
sciencemag.org/subscriptions

MEMBER BENEFITS

aaas.org/membercentral

AAAS BOARD OF DIRECTORS

CHAIR Steven Chu
PRESIDENT Claire M. Fraser
PRESIDENT-ELECT Susan G. Amara
TREASURER Carolyn N. Ainslie
CHIEF EXECUTIVE OFFICER
Sudip Parikh
BOARD Cynthia M. Beall
Rosina M. Bierbaum
Ann Bostrom
Stephen P.A. Fodor
S. James Gates, Jr.
Laura H. Greene
Kaye Husbands Fealing
Maria M. Klawe
Robert B. Millard
Alonda Nelson
William D. Provine

BOARD OF REVIEWING EDITORS (Statistics board members indicated with \$)

Adriano Aguzzi, U. Hospital Zürich
Takuzo Aida, U. of Tokyo
Leslie Aiello, Wenner-Gren Foundation
Judith Allen, U. of Manchester
Sebastian Amigorena, Institut Curie
James Analytis, U. of California, Berkeley
Paola Ariotti, Harvard U.
Johan Auwerx, EPFL
David Awschalom, U. of Chicago
Clare Baker, U. of Cambridge
Nenad Ban, ETH Zürich
Franz Bauer, Pontificia Universidad Católica de Chile
Ray H. Baughman, U. of Texas at Dallas
Peter Bearman, Columbia U.
Carlo Beenakker, Leiden U.
Yasmine Belkaid, NIAID, NIH
Philip Benfey, Duke U.
Gabriele Bergers, VIB
Bradley Bernstein, Mass. General Hospital
Alessandra Biffi, Harvard Med. School
Peer Bork, EMBL
Chris Bowler, Ecole Normale Supérieure
Ian Boyd, U. of St. Andrews
Emily Brodsky, U. of California, Santa Cruz
Ron Brookmeyer, U. of California, Los Angeles (\$) **Christian Büchel, UKE Hamburg**
Dennis Burton, Scripps Research
Carter Tribble Butts, U. of California, Irvine
György Buzsáki, New York U. School of Med.
Blanche Capel, Duke U.
Annamarie Carlton, U. of California, Irvine
Nick Chater, U. of Warwick
Zhijian Chen, UT Southwestern Med. Ctr.
Ib Chorkendorff, Denmark TU
James J. Collins, MIT
Robert Cook-Deegan, Arizona State U.
Alan Cowman, Walter & Eliza Hall Inst.
Carolyn Coyne, U. of Pittsburgh
Roberta Croce, VU Amsterdam
Jeff L. Dangl, U. of North Carolina
Tom Daniel, U. of Washington
Chiara Daraio, Caltech
Nicolas Dauphas, U. of Chicago
Frans de Waal, Emory U.
Claude Desplan, New York U.
Sandra Diaz, Universidad Nacional de Córdoba
Ulrike Diebold, TU Wien
Hong Ding, Inst. of Physics, CAS
Jennifer Dionne, Stanford U.
Dennis Discher, U. of Penn.
Gerald Dorn, Washington U. in St. Louis
Jennifer A. Doudna, U. of California, Berkeley
Bruce Dunn, U. of California, Los Angeles
William Dunphy, Caltech
Christopher Dye, U. of Oxford
Todd Ehlers, U. of Tübingen
Jennifer Eilseesoff, Johns Hopkins U.
Tim Elston, U. of North Carolina
Andrea Encalada, U. San Francisco de Quito
Nader Engheta, U. of Penn.
Karen Ersche, U. of Cambridge
Barry Everitt, U. of Cambridge
Vanessa Ezenwa, U. of Georgia
Michael Feuer, The George Washington U.
Toren Finkel, U. of Pittsburgh Med. Ctr.
Gwen Flowers, Simon Fraser U.
Peter Fratzl, Max Planck Inst. Potsdam
Elaine Fuchs, Rockefeller U.
Eileen Furlong, EMBL
Jay Gallagher, U. of Wisconsin
Daniel Geschwind, U. of California, Los Angeles
Karl-Heinz Glassmeier, TU Braunschweig
Ramon Gonzalez, U. of South Florida
Elizabeth Grove, U. of Chicago
Nicolas Gruber, ETH Zürich
Hua Guo, U. of New Mexico
Kip Guy, U. of Kentucky College of Pharmacy
Taekjip Ha, Johns Hopkins U.
Christian Haass, Ludwig Maximilians U.
Sharon Hammes-Schiffer, Yale U.
Wolf-Dietrich Hardt, ETH Zürich
Louise Harra, U. College London
Jian He, Clemson U.
Carl-Philipp Heisenberg, IST Austria
Ykä Helariutta, U. of Cambridge
Janet G. Hering, Eawag
Hans Hilgenkamp, U. of Twente
Kai-Uwe Hinrichs, U. of Bremen
Lora Hooper, UT Southwestern Med. Ctr.
Fred Hughson, Princeton U.
Randall Hulet, Rice U.
Auke Ijspeert, EPFL
Akiko Iwasaki, Yale U.
Stephen Jackson, USGS and U. of Arizona
Kai Johnson, EPFL
Peter Jonas, IST Austria
Matt Kaeblerlein, U. of Washington
William Kaelin Jr., Dana-Farber Cancer Inst.
Daniel Kammen, U. of California, Berkeley
V. Naray Kim, Seoul Nat. U.
Robert Kingston, Harvard Med. School
Nancy Knowlton, Smithsonian Institution
Etienne Koechlin, Ecole Normale Supérieure
Alexander L. Kolodkin, Johns Hopkins U.
Julija Krupic, U. of Cambridge
Thomas Langer, Max Planck Inst. Cologne
Mitchell A. Lazar, U. of Penn.
Ottoline Leyser, U. of Cambridge
Wendell Lim, U. of California, San Francisco
Jiangui Liu, Michigan State U.
Luis Liz-Marzán, CIC biomaGUNE
Jonathan Losos, Washington U. in St. Louis
Ke Lu, Chinese Acad. of Sciences
Christian Lüscher, U. of Geneva
Jean Lynch-Stieglitz, Georgia Inst. of Tech.
Fabienne Mackay, U. of Melbourne
Anne Magurran, U. of St. Andrews
Oscar Marín, King's College London
Charles Marshall, U. of California, Berkeley
Christopher Marx, U. of Idaho
Geraldine Masson, CNRS
C. Robertson McClung, Dartmouth College
Rodrigo Medellín, U. Nacional Autónoma de México
Graham Medley, London School of Hygiene & Tropical Med.
Jane Memmott, U. of Bristol
Baoxia Xi, U. of California, Berkeley
Edward Miguel, U. of California, Berkeley
Tom Misteli, NCI, NIH
Yasushi Miyashita, U. of Tokyo
Alison Motingser-Reif, NIEHS, NIH (\$) **Daniel Nettle, Newcastle U.**
Daniel Neumark, U. of California, Berkeley
Beatriz Noheida, U. of Groningen
Helga Nowotny, Vienna Science, Research & Tech. Fund
Rachel O'Reilly, U. of Birmingham
Harry Orr, U. of Minnesota
Pilar Ossorio, U. of Wisconsin
Andrew Oswald, U. of Warwick
Isabella Pagano, Istituto Nazionale di Astrofisica
Margaret Palmer, U. of Maryland
Elizabeth Levy Paluck, Princeton U.
Jane Parker, Max Planck Inst. Cologne
Giovanni Parmigiani, Dana-Farber Cancer Inst. (\$) **Samuel Pfaff, Salk Inst. for Biological Studies**
Julie Pfeiffer, UT Southwestern Med. Ctr.
Matthieu Piel, Institut Curie
Kathrin Plath, U. of California, Los Angeles
Martin Pollard, U. of Penn.
Katherine Pollard, U. of California, San Francisco
Elvira Poloczanska, Alfred-Wegener-Inst.
Julia Pongratz, Ludwig Maximilians U.
Philippe Poulin, CNRS
Jonathan Pritchard, Stanford U.
Félix A. Rey, Institut Pasteur
Trevor Robbins, U. of Cambridge
Joeri Rogelj, Imperial College London
Amy Rosenzweig, Northwestern U.
Mike Ryan, U. of Texas at Austin
Mitinori Saitou, Kyoto U.
Shimon Sakaguchi, Osaka U.
Miquel Salmeron, Lawrence Berkeley Nat. Lab
Nitin Samarth, Penn. State U.
Jürgen Sandkühner, Med. U. of Vienna
Alexander Schier, Harvard U.
Wolfram Schlenker, Columbia U.
Susannah Scott, U. of California, Santa Barbara
Rebecca Sear, London School of Hygiene & Tropical Med.
Vladimir Shalae, Purdue U.
Jie Shan, Cornell U.
Beth Shapiro, U. of California, Santa Cruz
Jay Shendure, U. of Washington
Steve Sherwood, U. of New South Wales
Brian Shoichet, U. of California, San Francisco
Robert Siliciano, Johns Hopkins U. School of Med.
Lucia Sivilotti, U. College London
Alison Smith, John Innes Centre
Richard Smith, U. of North Carolina (\$) **Mark Smyth, QIMR Berghofer**
Pam Soltis, U. of Florida
John Speakman, U. of Aberdeen
Tara Spire-Jones, U. of Edinburgh
Allan C. Spradling, Carnegie Institution for Science
V. S. Subrahmanian, Dartmouth College
Ira Tabas, Columbia U.
Sarah Teichmann, Wellcome Sanger Inst.
Rocio Titiunik, Princeton U.
Shubha Tole, Tata Inst. of Fundamental Research
Wim van der Putten, Netherlands Inst. of Ecology
Reinhold Veugeler, KU Leuven
Bert Vogelstein, Johns Hopkins U.
Kathleen Vohs, U. of Minnesota
David Wallace, Weizmann Inst. of Science
Jane-Ling Wang, U. of California, Davis (\$) **David Waxman, Fudan U.**
Jonathan Weissman, U. of California, San Francisco
Chris Winkle, U. of Missouri (\$) **Terrie Williams, U. of California, Santa Cruz**
Ian A. Wilson, Scripps Research (\$) **Yu Xie, Princeton U.**
Jan Zaenen, Leiden U.
Kenneth Zaret, U. of Penn. School of Med.
Jonathan Zehr, U. of California, Santa Cruz
Xiaowei Zhuang, Harvard U.
Maria Zuber, MIT

COVID-19 research in Africa

When two French doctors recently discussed the ease of conducting clinical research on coronavirus disease 2019 (COVID-19) in African nations, with an insinuation that ethical and safety standards for testing vaccines and treatments in these nations are lower than in other countries, anti-African research sentiments flared on social media and in news reports. Suggesting that clinical trial conduct is at a lower standard in Africa is unacceptable. Africa has innovated and implemented health solutions with high ethical regard for its people. In March, the Academy of Science of South Africa stressed the importance of research and development on COVID-19 in Africa as “key to the response to outbreaks of emerging and re-emerging pathogens.” The expertise and infrastructure of African clinical research sites, along with African nations’ engagement of their communities who wish to contribute to effective solutions, are ready to be leveraged to tackle this pandemic emergency.

So far, COVID-19, caused by severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2), has infected ~5 million people worldwide and caused ~328,000 deaths. Africa now has over 95,000 cases of infection across the continent and more than 3000 deaths, with Egypt (14,000) just behind South Africa (18,000) in reported infections. But Africa has long grappled with the morbidity and mortality of communicable diseases including endemic tuberculosis (TB), Ebola virus disease, and malaria. Sub-Saharan Africa has borne the brunt of the HIV/AIDS epidemic, with half of the annual global infections. Once again, Africa confronts a new pandemic and must help find solutions, both for the continent and for the global community.

Africa, a consumer of health products, has played a key role in developing new medical products. Driven by the need to improve local public health, highly collaborative partnerships have been established between international research teams that involve African scientists, clinicians, and community-based advocates. For example, at the turn of the millennium, HIV/AIDS was a death sentence throughout Africa; that is no longer the case. Today, many antiretroviral drugs used worldwide were tested in Sub-Saharan Africa and found to be lifesaving for people with advanced AIDS. The extensive involvement of African women in clinical HIV/AIDS research underpinned the establishment of new preventive tools.

Also developed in Africa were the guidelines for

treating AIDS and TB coinfection, cryptococcal meningitis infection, immune reconstitution syndrome, the timing of pediatric antiretroviral therapy initiation, and the evaluation of new therapeutic regimens for treating drug-resistant TB. Vaccines for malaria and Ebola virus disease were also tested across Africa. Recently, vaccine candidates for TB prevention were tested in East, West, Central, and Southern Africa, a necessary prerequisite before deployment across the continent.

The key components to Africa’s contributions in these endeavors have included robust community engagement with clinical research, the ethical conduct of this research, and ensuring that the research is regulated, monitored, and analyzed within Africa. These principles must also guide the partnerships needed to address COVID-19 effectively. For example, local communities should be considered partners in clinical research.

All research should have ethical and regulatory approval in-country. And the highest international and national standards of treatment and prevention for COVID-19 should be maintained.

Several countries in Africa are now preparing for clinical trials of COVID-19 therapies and vaccines with multiple international partners. Cameroon, Zambia, Zimbabwe, Uganda, and South Africa are applying for ethical and regulatory approval for the Chloroquine Repurposing to Health Workers for Novel Coronavirus Mitigation Trial (supported by Washington University School of Medi-

cine and the Bill and Melinda Gates Foundation). Nigeria, Tunisia, Egypt, and South Africa have signed up for the Solidarity Trial [supported by the World Health Organization (WHO) and partners] to rapidly determine whether treatment options slow COVID-19 progression or improve survival. National regulatory authorities and national ethics committees from across Africa have agreed to combine their expertise to expedite clinical trial review and approvals for new multinational preventive, diagnostic, and therapeutic interventions for COVID-19. Although each country is solely responsible for granting regulatory approval, this agreement was reached during a virtual meeting convened by the WHO on 1 April 2020 under the platform of the African Vaccines Regulatory Forum.

As with so many other diseases, COVID-19 trials will be carried out in Africa, under the highest ethical and safety standards. To exclude Africa would be a life-threatening mistake.

—Linda-Gail Bekker and Valerie Mizrahi*

Linda-Gail Bekker is deputy director at The Desmond Tutu HIV Centre, University of Cape Town, Cape Town, South Africa. linda-gail.bekker@hiv-research.org.za

Valerie Mizrahi is director at the Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town, South Africa. mizrahi@uct.ac.za

“To exclude Africa would be a life-threatening mistake.”

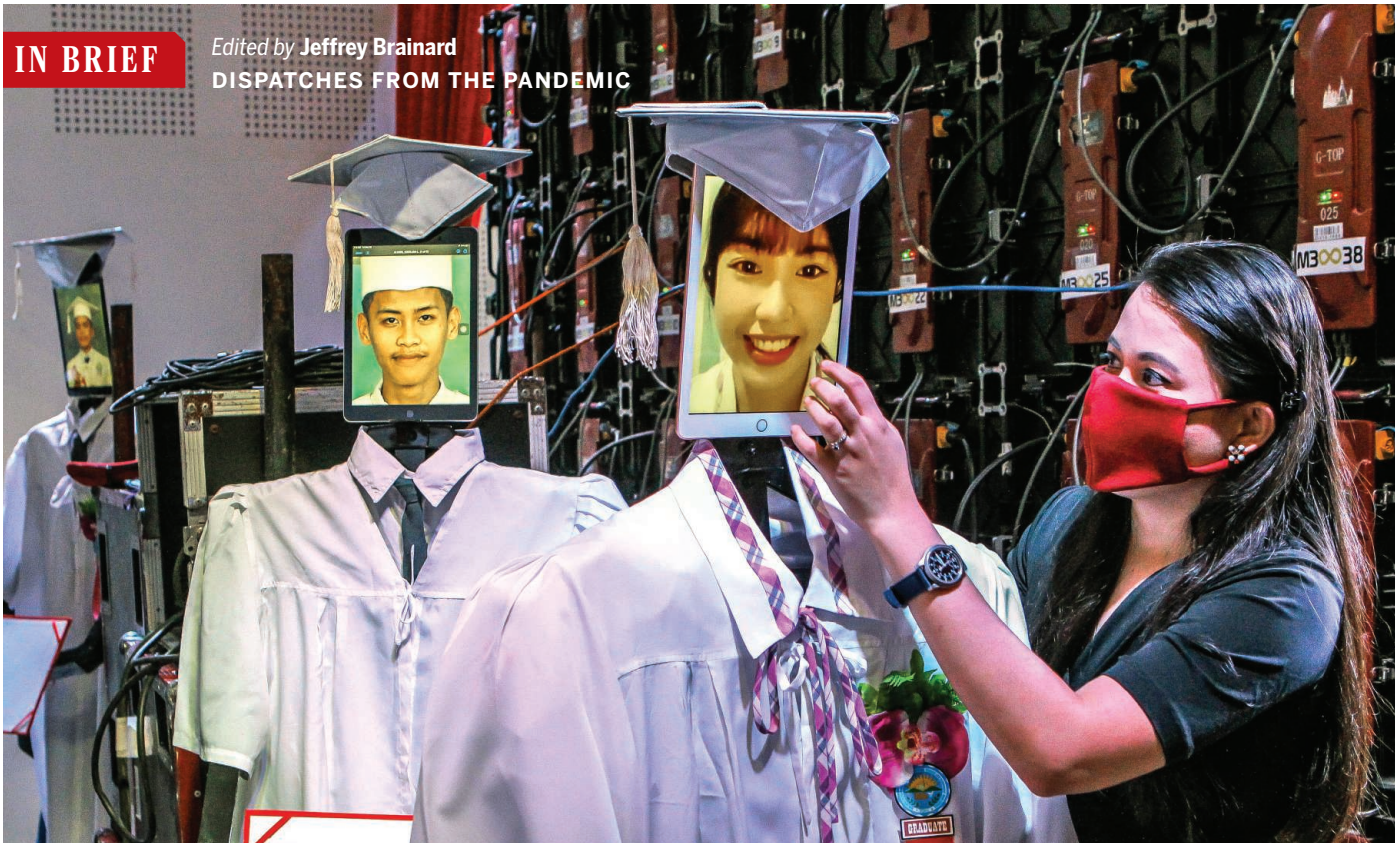
*Acknowledgments are in the supplementary materials (science.sciencemag.org/content/368/6494/919/suppl/DC1).

“It’s like the house was on fire.
[They] could not provide substantial help.”

Virologist Shao Yiming of China’s Center for Disease Control and Prevention, when asked why China barred entry to scientists from other countries to investigate COVID-19 early in the pandemic.

IN BRIEF

Edited by Jeffrey Brainard
DISPATCHES FROM THE PANDEMIC



A teacher prepares a tablet showing a student's image for a “cybergraduation” ceremony at a high school in Manila, Philippines, on 22 May, as social distancing continued.

U.S. goes big for Oxford vaccine

CLINICAL TRIALS | The United States last week placed a \$1.2 billion bet on a candidate COVID-19 vaccine, its biggest so far. Operation Warp Speed, the White House’s bid to quickly deliver a vaccine, is investing in a candidate manufactured by AstraZeneca, which would by October begin to provide the first of at least 300 million doses. Researchers at the University of Oxford developed the vaccine, which contains a harmless virus engineered to hold the gene for the spike surface protein of the coronavirus. It prevented COVID-19-like disease in a monkey experiment, although all of the animals became infected. An ongoing clinical trial in the United Kingdom should provide initial data on the vaccine’s safety and could even reveal signs of efficacy. Warp Speed’s selection process for vaccine candidates is not public,

and some scientists worry the rush for a vaccine, shortcutting the standard vaccine development process, is designed to boost President Donald Trump’s chance of re-election in November. In a separate COVID-19 vaccine development, Merck, a major pharmaceutical company, has joined the race, announcing collaborations with groups working on two different candidates.

States stop mixing test data

DIAGNOSTICS | Multiple states and the U.S. Centers for Disease Control and Prevention (CDC) were under fire last week for releasing coronavirus data that mixed tests for active and past infections. Scientists said that practice gave an inaccurate picture of the number of current COVID-19 cases and exaggerated the nation’s capacity for diagnosing active cases. Diagnostic tests for active

infections, usually based on nose swabs, identify the presence of genes from the virus that causes the disease. Other tests detect antibodies to the virus in blood samples, identifying people likely no longer infected. Officials in Georgia and Virginia said last week they would end the mixed tallies, which were reported by *The Atlantic* and several newspapers. CDC says it has been working to standardize reporting requirements across states and plans to disaggregate them by test type on its website in the coming weeks.

Dengue tied to Singapore closure

EPIDEMIOLOGY | Scientists in Singapore suspect the city’s COVID-19 lockdown may have caused a spike in dengue fever, a viral disease that causes intense pain and is sometimes fatal. The 6900 cases recorded since January are double that of the same period in 2019. Researchers say one reason

is that people staying home are more likely to be bitten by the mosquitoes that transmit the virus. The mosquito, *Aedes aegypti*, is more abundant in residential areas, which have more greenery and breeding sites, than in business districts. It bites during the day, when many people would otherwise be at school or work. Other tropical cities should be on the lookout for a surge as well, dengue specialists say. The good news is that social distancing and enhanced hygiene during COVID-19 have led to a sharp drop in infectious diseases not spread by mosquitoes, such as conjunctivitis and hand, foot, and mouth disease.

Nobel awardees decry grant halt

FUNDING | Seventy-seven U.S. scientists who have won a Nobel Prize last week asked federal officials to “act urgently” to review a controversial decision by the U.S. National Institutes of Health (NIH) to terminate a grant that supported research into bat coronaviruses in China. NIH’s explanation for killing the grant was “preposterous,” the laureates wrote in a letter to NIH Director Francis Collins and Health and Human Services Secretary Alex Azar. “We believe that this action sets a dangerous precedent by interfering in the conduct of science.” Thirty-one scientific societies also wrote last week to Collins asking that NIH’s action “be immediately reconsidered.” On 24 April, NIH informed the nonprofit EcoHealth Alliance that it was ending a grant, first awarded in 2014 and renewed in 2019, because it no longer “align[ed] ... with agency priorities.” The move came 1 week after President Donald Trump vowed to end the grant quickly, and after conservative U.S. politicians and media suggested—without evidence—that the coronavirus causing the pandemic escaped from a laboratory in Wuhan, China, that employs Chinese virologist Shi Zhengli, who had received funding from the grant.

AAAS to meet online in 2021

SCIENTIFIC SOCIETIES | AAAS, which publishes *Science*, announced last week its 2021 meeting will be online only because of the COVID-19 pandemic. The meeting was to have been held in Phoenix. AAAS leaders said that deciding now not to meet in person will allow time to plan an online-only program. AAAS plans to announce changes to the meeting program and format for scientific symposia in early June.

SCIENCEMAG.ORG/TAGS/CORONAVIRUS

Read additional *Science* coverage of the pandemic.



One of two “dueling dinosaur” fossils that led to a lawsuit. The owner plans to sell them to a museum.

PALEONTOLOGY

Court settles Montana fossil fight

A ruling by the Montana Supreme Court last week is expected to resolve a legal saga that threatened to upend fossil hunting in the dinosaur-rich state. The court decided that a valuable trove of fossils discovered at a ranch in 2006 is not legally the same as minerals such as gold or copper. The ruling nullifies claims by previous owners of the land that mineral rights they retained on the property gave them partial ownership of the fossils, which include a complete *Tyrannosaurus rex* skeleton and a pair of entwined dinosaurs that might have been locked in combat. The outcome is a win for scientists who had warned that tying fossils to mineral rights could make it harder to get permission to excavate and could throw into doubt who owns fossils already on display.

Users pan PubMed makeover

PUBLICATIONS | Scientists are fuming over a sweeping redesign of PubMed, a massive database of biomedical literature maintained by the U.S. National Center for Biotechnology Information. Some took to social media to excoriate the new styling and layout, changes in the display of search results, and supposedly enhanced search algorithms unveiled last week. In response, PubMed’s managers restored the search tool’s legacy version on a “short-term” basis.

Bill would expand NSF

FUNDING | Bipartisan legislation introduced last week in both houses of the U.S. Congress would expand the National Science Foundation (NSF) with a new technology directorate. A sum of \$100 billion would be spent over 5 years on both basic research and precommercial development in 10 technology fields deemed essential to keep the country

ahead of China and the rest of the world. The budget for the new directorate would grow to four times the size of the current \$8 billion NSF, which would be renamed the National Science and Technology Foundation.

NASA honors ‘mother of Hubble’

ASTRONOMY | A future NASA space telescope has a new name. The Wide Field Infrared Survey Telescope (WFIRST) is now the Nancy Grace Roman Space Telescope, after a pioneering agency astronomer sometimes called the mother of the Hubble Space Telescope, the agency said on 20 May. Roman joined the fledgling space agency in 1959 as its first chief astronomer. Realizing that orbiting telescopes could observe around the clock, and more clearly, she helped persuade NASA to launch a series of them, including the Hubble telescope. The \$3.2 billion Roman Space Telescope, due for launch in 2025, will in a single view take in 100 times more sky than Hubble.



A recovered COVID-19 patient donates plasma at a clinic in Bangkok.

COVID-19

Scientists put survivors' blood plasma to the test

Randomized trials are underway after small studies suggest transfusions can save lives

By Kai Kupferschmidt

On 13 March, with the COVID-19 pandemic exploding and drugs elusive, Arturo Casadevall published what he considers “maybe the most important paper” of his long career. In *The Journal of Clinical Investigation*, the infectious disease specialist at Johns Hopkins University and Liise-anne Pirofski of Albert Einstein College of Medicine argued that one effective treatment might already be at hand: the blood plasma of people who have recovered from the disease, rich in antibodies against the virus. The strategy seems to have worked in other infections, the duo pointed out, and the infrastructure for collecting and administering plasma exists. The risks are known and comparatively low. “We recommend that institutions ... begin preparations as soon as possible,” they wrote. “Time is of the essence.”

Ten weeks later, more than 16,000 patients at hundreds of U.S. hospitals have received the experimental therapy, and hope that it works could soon give way to evidence. A study of patients treated with serum at Mount Sinai Hospital in New York City, published as a preprint on 22 May, offers hints it may, as do other small studies elsewhere. But randomized controlled clinical trials (RCTs) that will give more definitive answers are still underway.

Blood or plasma from recovered patients has been tried as a therapy since at least the Spanish flu of 1918; reports from that pandemic suggest it helped. It has also been used to fight measles, severe acute respiratory syndrome, and lesser known diseases such as Argentine hemorrhagic fever. In a 1970s study of 188 patients with that disease, only 1% of plasma recipients died, versus 16.5% in a control group. “I think that it has a high likelihood [of working] based on history,” Casadevall says.

But some examples are less encouraging. In a study of 84 Ebola patients in Guinea in 2015, doctors did not see a benefit from convalescent plasma. (It's not clear why; perhaps the plasma just didn't contain that many potent antibodies.) And the treatment carries risks: Transfusions can transmit blood-borne pathogens, and in rare cases lead to conditions such as transfusion-related acute lung injury (TRALI), in which transferred antibodies damage pulmonary blood vessels, or transfusion-associated circulatory overload (TACO), when the patient's body doesn't adapt to the added blood volume, which can be up to half a liter. Both can lead to difficulty breathing and death.

Chinese doctors began to experiment with convalescent plasma in COVID-19 patients in January. In an April study published in the

Proceedings of the National Academy of Sciences, they reported that 10 out of 10 plasma recipients improved, whereas three out of 10 “matched controls”—people with the same characteristics who didn't get the treatment—died. Further small studies from China, Italy, and elsewhere also looked promising.

The Mount Sinai case series is the largest so far. The researchers didn't take the time to set up a randomized trial because “time was a luxury we didn't have in NYC,” Mount Sinai virologist Nicole Bouvier, one of the study's authors, wrote in an email. But because plasma was scarce in the early phase of the pandemic, plenty of patients did not get it, enabling the team to set up a matched control

trial comparing 39 severe COVID-19 patients who received plasma with four times as many patients who didn't.

The difference in mortality—12.8% in the plasma group and 24.4% in the control group—was not statistically significant, but when the team compared the patients' supplemental oxygen needs after transfusion, those on plasma did significantly better. The matched-control design can lead to biases, Bouvier concedes, but “at least we showed that there is some benefit to convalescent plasma.” The same researchers are already working on a study looking at data from about 275 treated patients, she says.

Science's COVID-19 coverage is supported by the Pulitzer Center.

Nahid Bhadelia, an infectious disease physician at Boston University, agrees the data so far are promising, but says only an RCT can give a final answer. Such trials are now underway in Germany, the United Kingdom, and the United States; results are expected in the months ahead.

Researchers have already collected more data on complications, and they seem to be rare. A U.S. paper looking only at the therapy's safety in the first 5000 patients found 36 severe adverse events, including TRALI and TACO cases, but some may have been the result of COVID-19 itself. Only two events were "definitely related" to the transfusion, according to the treating physician; 23 others were deemed "possibly" or "probably" related. "I wouldn't say the [safety] concerns have been put to rest, but they have been given a nap," says one of the authors, Michael Joyner of the Mayo Clinic.

Convalescent serum could also help prevent infection in those at high risk. In a trial coordinated by Johns Hopkins, 150 health care workers exposed to COVID-19 while not wearing proper protection will receive either convalescent serum or serum collected last year. Researchers will compare how many people in each group develop disease.

If convalescent plasma is shown to work, much more of it may be needed, and supply could become a challenge, Bhadelia says. One plasma donation—the volume depends on the donor's weight but it's usually between 690 and 880 milliliters in the United States—is enough for just one or two patients, and the donor's blood type needs to match the recipient's. But recovered patients might be able to donate plasma multiple times. In New York City, there is now more than enough to go around, in part because thousands of members of the hard-hit Orthodox Jewish community have donated.

Consistency is also an issue. The mix and concentration of antibodies differs from one donor to the next, which "is one of the unfortunate reasons why the clinical evidence generated around convalescent plasma has remained rather shallow," says Thomas Kreil, head of pathogen safety at Japanese pharma company Takeda. Together with several partners, Takeda is working to produce a product called hyperimmune globulin, for which the blood of hundreds of recovered patients is pooled and the antibodies concentrated about 10-fold. Hyperimmune globulin has a longer shelf life than plasma, and its higher concentration would allow doctors to give more antibodies to patients without the risk of TACO. An efficacy trial, funded by the U.S. National Institutes of Health, could start this summer. ■

COVID-19

Doctors race to understand inflammatory condition in kids

Studies are launching to pinpoint children most at risk

By Jennifer Couzin-Frankel

Three children at one London hospital in mid-April, followed the next day by three at another—for Elizabeth Whittaker, a pediatric infectious disease doctor at Imperial College London, those first cases raised an alarm. The youngsters had fevers, rashes, stomach pain, and, in some cases, heart problems, along with blood markers that characterize COVID-19 in adults, including one associated with clotting. But in most, nasal swabs failed to reveal any virus.

"I don't understand—they look like they have coronavirus," Whittaker recalls thinking. Doctors nonetheless suspected a link. Within days, a survey turned up 19 additional cases across England, and an alert on 27 April asked doctors to be on the lookout for such symptoms in children. Soon after, dozens more cases surfaced in New York along with smaller clusters elsewhere, bolstering a connection to the pandemic. Reports of children on life support and some deaths put parents on edge—and were especially disheartening after earlier signs that COVID-19 largely spares children from serious illness.

It is another surprise from a virus that has proffered many, and projects worldwide

are gearing up to study it. They are combing the blood and sequencing the genomes of patients—and the virus, if it can be isolated from them—to search for clues to what makes some children susceptible and how to head off the worst symptoms. There's hope that what's learned from young patients might help the many adults in whom COVID-19 also triggers a grievous overreaction of the immune system.

In some respects, "It's absolutely not shocking" to see this, says Rae Yeung, a rheumatologist and immunologist at the Hospital for Sick Children in Toronto, whose center treated about 20 children with similar symptoms over the past 3 weeks. Many pathogens occasionally trigger a similar hyperactive immune response in children, known as Kawasaki disease. Its symptoms vary but include rash, fever, and inflammation in medium-size blood vessels. Children can suffer heart complications. In rare cases, blood pressure plummets and shock sets in.

Doctors disagree on whether the variant linked to COVID-19 is Kawasaki disease or something new, with some experts calling it multisystem inflammatory syndrome in children. But as with Kawasaki disease, most patients recover with treatment, including steroids and immunoglobulins, which calm the immune system.



A girl in New Delhi gets a nasal swab to test for the new coronavirus.

In linking the inflammatory syndrome to COVID-19, “We’re going on more than just a hunch,” says Jesse Papenburg, a pediatric infectious disease specialist at Montreal Children’s Hospital, in a city that’s seen about 25 children with the condition. Kawasaki disease is rare, ordinarily affecting just one to three in every 10,000 children in Western countries, though it’s more common in children with Asian ancestry. The spikes recorded so far, in COVID-19 hot spots like northern Italy and New York City, track the novel coronavirus’ march around the world. And although a minority of these children test positive for SARS-CoV-2, a study published in *The Lancet* by a team in Bergamo, Italy, reported that eight of 10 children with the Kawasaki-like illness had antibodies to the virus, indicating they had been infected. Positive antibody tests have been reported in sick children elsewhere, too.

“It was obvious that there was a link,” says Lorenzo D’Antiga, a pediatrician at the Papa Giovanni XXIII Hospital who led the study. The new coronavirus can elicit a powerful immune response, which he thinks may explain why shock and a massive immune reaction called a cytokine storm are more common in the COVID-19-linked cases than in textbook Kawasaki disease. And a time lag between infection and the Kawasaki-like illness could explain why many of the affected children show no evidence of the virus. The immune system’s overreaction may unfold over weeks, though virus could also be hiding somewhere in the body.

“There’s clearly some underlying genetic component” that puts a small number of children at risk, says Tom Maniatis, founding director of Columbia University’s Precision Medicine Initiative. New York state is investigating at least 157 cases, and Maniatis is also CEO of the New York Genome Center, which is pursuing whole-genome sequencing of affected children and their parents, as well as sequencing the virus found in children, with family consent. Finding genes that heighten risk of the illness or of developing a severe case could point to better treatments or help identify children who may take a sudden turn for the worse.

Genetics may also help explain a puzzle: why the illness hasn’t been reported in Asian countries, even though Kawasaki disease is far more common in children with Asian ancestry. The virus’ own genetics may be important; an analysis last month indicated that the predominant viral variant in New York was brought by travelers from Europe. It’s also possible that the Kawasaki-like illness is so rare that it only shows up in COVID-19 hotbeds. “The areas that have been hardest

hit by coronavirus are the areas reporting this syndrome now,” says Alan Schroeder, a critical care physician at Lucile Packard Children’s Hospital at Stanford University, which has seen one potentially affected child, a 6-month-old baby, who recovered quickly.

Yeung is also pursuing ways to flag children with COVID-19 who are at risk of this complication. She co-leads an international consortium that’s banking blood from affected children, both before and after treatment, and screening for various markers, including the cytokine molecules that indicate a revved-up immune system. The group is also searching for gene variants known to predict poor outcomes in Kawasaki disease. “There’s also core COVID stuff that needs to be measured,” Yeung says, such as markers of heart function and levels of D-dimer, a protein fragment in the blood that indicates a tendency toward clotting and that surges in many sick adults.

A European Union Horizon 2020 project called DIAMONDS, originally designed to improve diagnosis of pathogens in children with fevers, is recruiting children across Europe with the Kawasaki-like complication, along with those who have run of the mill COVID-19 symptoms. Scientists will study blood for pathogens—not just SARS-CoV-2—and the behavior of immune cells such as T cells and B cells.

“We have to do a deep dive into the immunology of those patients,” says Elie Haddad, a pediatric immunologist and scientist at Sainte-Justine University Hospital Center who, with Yeung and Susanne Benseler at Alberta Children’s Hospital, is leading a Canadian research effort on the new syndrome. These deep dives may also clarify the immune system chaos seen in many sick adults. Children are “cleaner,” Haddad points out—they’re less likely to have other health burdens, such as diabetes or high blood pressure, that can make it harder to tease out the virus’ impact on the immune system.

Last week, young adults with possible cases of the condition were identified, suggesting it may not be limited to children. A global effort studying COVID-19 in adults, called the International Severe Acute Respiratory and Emerging Infection Consortium, will look at adults’ clinical data and blood samples, Whittaker says, “to see, is this a uniquely pediatric problem?”

Eager as they are to understand this new face of the pandemic, doctors want to avoid overstating the hazards. “We need to identify early and we need to intervene early” in treating these children, Yeung says. But she also urges calm. “The kids we’re seeing so far,” she stresses, “they respond to the treatments we’re giving.” ■



By Jeffrey Brainard

Timothy Sheahan, a virologist studying COVID-19, wishes he could keep pace with the growing torrent of new scientific papers related to the pandemic. But there have just been too many—more than 5000 papers a week. “I’m not keeping up,” says Sheahan, who works at the University of North Carolina, Chapel Hill. “It’s impossible.”

A loose-knit army of data scientists and software developers is pressing hard to change that. They are creating digital collections of papers and building search tools powered by artificial intelligence (AI) that could help researchers quickly find the information they seek. The urgency is growing: The COVID-19 literature has grown to more than 31,000 papers since January and by one estimate is on pace to hit more than 52,000 by mid-June—among the biggest explosions of scientific literature ever.

The volume of information “is like what you would get in a medical conference that used to happen yearly. Now, that’s happening daily,” says Sherry Chou, a neurologist at the University of Pittsburgh Medical Center who is studying COVID-19’s neurologic effects.

“People don’t have time to read through entire articles and figure out what is the value added ... and what are the limitations,” says Kate Grabowski, an epidemiologist at Johns Hopkins University’s School of Medicine who leads an effort to create a curated set of pandemic papers.

It’s not clear, however, whether the emerging efforts will tame the tsunami.

ILLUSTRATION: SARA GIRONI CARNEVALE

New tools aim to tame pandemic paper tsunami

Researchers harness AI and old-fashioned expertise to glean findings, but face challenges



Despite a global effort to persuade publishers to make all papers relevant to COVID-19 immediately free, as many as 20% of new papers are still behind paywalls, a recent study found, off-limits to some readers and AI analysis. Some of the new search tools, meanwhile, aren't very user-friendly or are little known. And many researchers are skeptical that the tools will tell them what they really want to know: What is the work's quality? "People tend to oversell and put up papers with data that do not support their conclusions," Sheahan says. "It's a mess."

One line of work got a boost on 16 March, when the White House Office of Science and Technology Policy announced the launch of the COVID-19 Open Research Dataset (CORD-19), a trove that now includes more than 128,000 peer-reviewed articles and preprints, including studies of virology and coronaviruses dating back decades. To create the archive, some of the largest groups active in machine learning—including Google, the Chan Zuckerberg Initiative, and the Allen Institute for AI—collaborated with the National Institutes of Health and others to use search methods, such as natural language processing, to scan the scientific literature for relevant terms. The team also converted PDF files into a form readable by machine learning algorithms so other researchers could analyze the papers.

CORD-19's creation was "amazing work," says Giovanni Colavizza, a bibliometrics researcher at the University of Amsterdam. But analyses he and colleagues conducted have found potential shortcomings. For example, as of 17 April, about 75% of the pa-

pers don't mention search terms used by CORD-19's creators, such as "coronavirus," in their titles, abstracts, or keywords, the researchers reported in a preprint posted on bioRxiv. That means these articles might only be tangentially related to COVID-19, he says. What's more, fewer than half the papers provided full text, necessary for comprehensive data mining by AI programs.

A growing number of papers are also not freely available to human readers. In response to calls from major science funders, including some governments, most major publishers have pledged to make free their COVID-19-related papers. But the number of paywalled publications is growing faster than the free ones, according to a study led by Nicolas Robinson-Garcia of the Delft University of Technology and posted as a preprint on 26 April on bioRxiv. By 1 June, nearly half of all COVID-19 papers could be behind paywalls, the researchers estimate, which also limits data mining and AI-enhanced searching.

Despite these limitations, many teams are turning to advanced computational tools to mine databases such as CORD-19. Data scientists, for example, have launched more than 1500 projects in response to a White House call to build tools that use CORD-19 to help answer 10 high-priority, pandemic-related research questions identified by the U.S. National Academy of Sciences and the World Health Organization.

One fruit of these efforts, which are listed on the Kaggle online hub, is an "AI-powered literature review." It used algorithms to harvest data points from more than 830 papers

in CORD-19 on 17 topics, and presents a web page for each topic that displays data tables and links to more information. But the algorithms don't always extract the data correctly, so medical students and other volunteers idled by the pandemic have been manually checking the tool's accuracy.

Another challenge is making some tools user friendly. A team at the Allen Institute for AI recently unveiled SciSight, which helps those searching the CORD-19 database by automatically suggesting similar papers and drawing browsable maps of related papers.

Grabowski's team at Johns Hopkins decided to emphasize human judgment over automated approaches. To create their 2019 Novel Coronavirus Research Compendium, which debuted on 17 April, more than 50 are combing through the literature. So far, they have selected and summarized more than 120 papers on eight topics, including vaccines and treatments. The team excluded most of the articles it examined because they only contained commentaries, protocols, poor-quality models, or no original findings, Grabowski says. The effort is focused on studies in humans, and the intended readers include health care workers and policymakers, Grabowski says. "We are trying to fill a void that we saw. ... [T]here is just so much information, but a lot of the studies are not conducted very well."

It's too soon to measure the quality of pandemic papers based on citations or retractions, specialists say. But if quality is suffering, it's not because, as often feared, large numbers of studies are bypassing peer review and appearing only as preprints. Preprints only make up a minority of the COVID-19 gusher, according to Robinson-Garcia's team. As of 14 April, some 80% of the more than 11,000 COVID-19 manuscripts it examined had appeared in refereed journals. (Some were preprints originally.)

Despite the many efforts to help them navigate the literature, Sheahan and other scientists say they have not heard of new tools released in recent weeks or have had little time to try them. And persuading researchers to adopt them could be difficult, says Jevin West, a data scientist at the University of Washington, Seattle. "It's going to take some time to get people to change their habits," he says. Making that pitch during a pandemic is "like going into an emergency room and giving the doctors a different scalpel and saying: 'This is actually better:'"

In the meantime, many researchers say they are falling back on time-tested ways to keep up with new results, including reading bulletins from scientific societies and a few leading journals, as well as relying on word of mouth—including tweets—from trusted colleagues. ■

CLIMATE CHANGE

UV radiation blamed in ancient mass extinction

Malformed spores suggest powerful storms drove ozone loss as climate warmed

By Paul Voosen

The end of the Devonian period, 359 million years ago, was an eventful time: Fish were inching out of the ocean, and fernlike forests were advancing on land. The world was recovering from a mass extinction 12 million years earlier, but the climate was still chaotic, swinging between hothouse conditions and freezes so deep that glaciers formed in the tropics. And then, just as the planet was warming from one of these ice ages, another extinction struck, seemingly without reason. Now, spores from fernlike plants, preserved in ancient lake sediments from eastern Greenland, suggest a culprit: The planet's protective ozone layer was suddenly stripped away, exposing surface life to a blast of mutation-causing ultraviolet (UV) radiation.

Just as the extinction set in, the spores became misshapen and dark, indicating DNA damage, John Marshall, a palynologist at the University of Southampton, and his co-authors say in a paper published this week in *Science Advances*. It's evidence, he says, that "all of the ozone protection is gone."

Scientists have long believed—at least before humanity became a force for extinction—that there were just two ways to wipe out life on Earth: an asteroid strike or massive volcanic eruptions. But 2 years ago, researchers found evidence that in Earth's worst extinction—the end-Permian, 252 million years ago—volcanoes lofted Siberian salt deposits into the stratosphere, where they might have fed chemical reactions that oblit-

erated the ozone layer and sterilized whole forests. Now, spores from the end-Devonian make a compelling case that, even without eruptions, a warming climate can deplete the ozone layer, says Lauren Sallan, a paleobiologist at the University of Pennsylvania. "Because the evidence is so strong, it will make people rethink other mass extinction events."

The end-Devonian die-off has long sat in the shadow of the Late Devonian extinction 12 million years earlier, one of the planet's largest. Likely driven by volcanoes that emitted gases that drastically cooled and warmed the planet, it killed most corals and many shelled sea creatures. But 10 years ago, work by Sallan and others revealed the end-Devonian was mighty in its own right, wiping out many plants and vertebrates, including most tetrapods, the four-limbed fish that had begun to evolve fingers and toes. Only the five-toed tetrapods survived. "It resets our own evolution," Marshall says. "All these archaic lineages, it kicked them out of the frame."

What the end-Devonian lacked was a cause. There was no evidence for volcanism or a giant impact, but one alluring clue was seen in the rapid formation and disappearance of rock deposits associated with glaciers, Sallan says. "Something was really screwed up with climate at that time."

Over the past 3 decades, Marshall has explored rocks surviving from this time in eastern Greenland. At the time, this terrain lay far from the arctic, at lower latitudes, locked in the arid interior of a landmass called the Old Red Sandstone Continent. As the climate warmed after the Devonian's last ice

An "overshooting" thunderstorm can inject water into the stratosphere, where it may destroy ozone.

age, lakes formed and filled with sediment that slowly turned to mudstone, recording conditions before and during the extinction. In 2017, Marshall exhumed the perfect mudstone in a 6-meter-long drilled core.

It captures a startling transformation: Healthy fossilized spores, coated in distinctive symmetrical spikes, suddenly grow misshapen, their spikes dilapidated and uneven. Spores are a common fossil because of their armored coat, but they are vulnerable to UV radiation, much like humans; spores can even develop a "tan" in response to UV. The damage Marshall saw is consistent with such exposure, says Jeffrey Benca, an experimental paleobotanist who has linked such damage to the end-Permian extinction. "What they propose seems quite plausible," he says.

Marshall argues that the warming climate drove more powerful summer thunderstorms, which could have injected an ozone-depleting mix of water and salts into the stratosphere. As UV rays killed off forests, nutrient runoff into the sea could have caused blooms of plankton and algae, which would have produced more ozone-destroying salts in a runaway feedback. "It looks like it might be a perfect storm," he says.

Marshall's scenario could explain not just the extinction, but also the many natural gas deposits dating from the period, says Sarah Carmichael, a geochemist at Appalachian State University. They formed from decaying organic matter, but no one has explained the needed surge in plankton growth. Nutrient runoff from dead forests could have fertilized the marine life.

It's also a portent of what could happen in today's warming world, where more powerful thunderstorms sometimes "overshoot" the troposphere and inject moisture into the dry, cold stratosphere. When combined with aerosol particles and chlorine molecules, the moisture may eat away ozone (*Science*, 26 April 2019, p. 322).

But atmospheric scientists can barely agree on whether these ozone depletions are happening now, let alone hundreds of millions of years ago. More overshoots occur now than expected, but whether they are spurring damaging reactions is not yet clear. Elliot Atlas, an atmospheric chemist at the University of Miami who studies this dynamic, is skeptical of Marshall's theory. It needs much more rigorous testing in models, he says. "Is it impossible? I can't say that."

Carmichael, for her part, would like to see evidence beyond the pollen grains that UV drove the extinction. "I'm wary of saying UV radiation is the reason," she says. "But I think it's a reason." ■

Core progress in AI has stalled in some fields

When tuned up, old algorithms can match the abilities of their successors

By **Matthew Hutson**

Artificial intelligence (AI) just seems to get smarter and smarter. Each iPhone learns your face, voice, and habits better than the last, and the threats AI poses to privacy and jobs continue to grow. The surge reflects faster chips, more data, and better algorithms. But some of the improvement comes from tweaks rather than the core innovations their inventors claim—and some of the gains may not exist at all, says Davis Blalock, a computer science graduate student at the Massachusetts Institute of Technology (MIT). Blalock and his colleagues compared dozens of approaches to improving neural networks—software architectures that loosely mimic the brain. “Fifty papers in,” he says, “it became clear that it wasn’t obvious what the state of the art even was.”

The researchers evaluated 81 pruning algorithms, programs that make neural networks more efficient by trimming unneeded connections. All claimed superiority in slightly different ways. But they were rarely compared properly—and when the researchers tried to evaluate them side by side, there was no clear evidence of performance improvements over a 10-year period. The result, presented in March at the Machine Learning and Systems conference, surprised Blalock’s Ph.D. adviser, MIT computer scientist John Guttag, who says the uneven comparisons themselves may explain the stagnation. “It’s the old saw, right?” Guttag said. “If you can’t measure something, it’s hard to make it better.”

Researchers are waking up to the signs of shaky progress across many subfields of AI. A 2019 meta-analysis of information retrieval algorithms used in search engines concluded the “high-water mark ... was actually set in 2009.” Another study in 2019 reproduced seven neural network recommendation systems, of the kind used by media streaming services. It found that six failed to outperform much simpler, non-neural algorithms developed years before, when the earlier techniques were fine-tuned, revealing “phantom progress” in the field. In another paper posted on arXiv in March, Kevin Musgrave, a computer scientist at Cor-

nell University, took a look at loss functions, the part of an algorithm that mathematically specifies its objective. Musgrave compared a dozen of them on equal footing, in a task involving image retrieval, and found that, contrary to their developers’ claims, accuracy had not improved since 2006 (see chart, below). “There’s always been these waves of hype,” Musgrave says.

Gains in machine-learning algorithms can come from fundamental changes in their architecture, loss function, or optimization strategy—how they use feedback to improve. But subtle tweaks to any of these can also boost performance, says Zico Kolter, a computer scientist at Carnegie Mellon University who studies image-recognition models trained to be immune to “adversarial at-

Other major algorithmic advances also seem to have stood the test of time. A big breakthrough came in 1997 with an architecture called long short-term memory (LSTM), used in language translation. When properly trained, LSTMs matched the performance of supposedly more advanced architectures developed 2 decades later. Another machine-learning breakthrough came in 2014 with generative adversarial networks (GANs), which pair networks in a create-and-critique cycle to sharpen their ability to produce images, for example. A 2018 paper reported that with enough computation, the original GAN method matches the abilities of methods from later years.

Kolter says researchers are more motivated to produce a new algorithm and tweak it until it’s state-of-the-art than to tune an existing one. The latter can appear less novel, he notes, making it “much harder to get a paper from.”

Guttag says there’s also a disincentive for inventors of an algorithm to thoroughly compare its performance with others—only to find that their breakthrough is not what they thought it was. “There’s a risk to comparing too carefully.” It’s also hard work: AI researchers use different data sets, tuning methods, performance metrics, and baselines. “It’s just not really feasible to do all the apples-to-apples comparisons.”

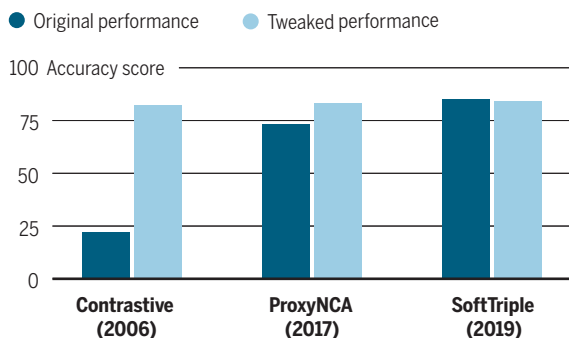
Some of the overstated performance claims can be chalked up to the explosive growth of the field, where papers outnumber experienced reviewers. “A lot of this seems to be growing pains,” Blalock says. He urges reviewers to insist on better comparisons to benchmarks and says better tools will help. Earlier this year, Blalock’s co-author, MIT researcher Jose Gonzalez Ortiz, released software called ShrinkBench that makes it easier to compare pruning algorithms.

Researchers point out that even if new methods aren’t fundamentally better than old ones, the tweaks they implement can be applied to their forebears. And every once in a while, a new algorithm will be an actual breakthrough. “It’s almost like a venture capital portfolio,” Blalock says, “where some of the businesses are not really working, but some are working spectacularly well.” ■

Matthew Hutson is a journalist in New York City.

Old dogs, new tricks

After modest tweaks, old image-retrieval algorithms perform as well as new ones, suggesting little actual innovation.



tacks” by a hacker. An early adversarial training method known as projected gradient descent (PGD), in which a model is simply trained on both real and deceptive examples, seemed to have been surpassed by more complex methods. But in a February arXiv paper, Kolter and his colleagues found that all of the methods performed about the same when a simple trick was used to enhance them.

“That was very surprising, that this hadn’t been discovered before,” says Leslie Rice, Kolter’s Ph.D. student. Kolter says his findings suggest innovations such as PGD are hard to come by, and are rarely improved in a substantial way. “It’s pretty clear that PGD is actually just the right algorithm,” he says. “It’s the obvious thing, and people want to find overly complex solutions.”

ANIMAL WELFARE

Sick chinchillas languish at farms that supply researchers

Years of animal welfare violations documented by USDA

By Meredith Wadman

Biomedical researchers rely on chinchillas—docile South American rodents with ears strikingly similar to those of humans—for studies of ear infections and hearing loss. But the two main U.S. chinchilla suppliers to research labs have for years violated the U.S. Animal Welfare Act, according to the U.S. Department of Agriculture (USDA), which enforces that law. Until the coronavirus pandemic descended, the two suppliers sent hundreds of animals to U.S. labs.

Those suppliers failed to identify and treat sick and injured animals, kept them in filthy barns and excrement-laden enclosures, and failed to clear dead animals, according to USDA inspection reports recently restored to full public view. One sup-

Animal Welfare Institute (AWI), which advocates for lab animals.

AALAS President Tracy Parker says its Buyers Guide lists only USDA-licensed research animal suppliers, but that the organization does not check compliance with the Animal Welfare Act. That's up to researchers and institutions, she says. "It's very troubling to me that it looks like research facilities that are continuing to utilize chinchillas haven't pressured [MCR] to improve. ... I would not allow animals from this facility into my program."

In a 2019 court filing, MCR proprietor Dan Moulton denied a long list of violations documented in a formal complaint that USDA filed in November 2018. He says he complied with the law, doing "very well for many years," until his USDA inspectors changed about 5 years ago. "The Animal

recent papers in journals including *Cell* and *Science Translational Medicine*.

USDA's complaint against MCR documented 85 animals that failed to receive veterinary care between September 2014 and May 2017. It described an excrement-ridden, understaffed barn housing lethargic animals with weeping, open wounds; putrid abscesses; and, most commonly, infected eyes. In seven out of eight inspections after the complaint was filed, the agency found continuing violations, including tight collars overlying open sores.

Filing a complaint is the most serious enforcement action USDA can take. Complaints can lead to a hearing at which a facility's license may be suspended or revoked. USDA says because of the pandemic, a new date has not been set for the postponed April hearing.

Moulton, a lawyer, told *Science* that he, his wife, and one 25-hour-per-week worker maintain 546 animals in a 771-square-meter barn. He adds that he supplies only healthy animals to researchers. "If the eyes are not bright, you don't send them. If the animals have got loose stool, you don't send them."

Sanford Feldman, director of comparative medicine at the University of Virginia, says he has consulted with Moulton in recent years as Moulton tried to rid his facility of endemic *Streptococcus zooepidemicus*, a pathogen that causes conjunctivitis, abscesses, and other problems. Feldman reviewed MCR's recent USDA inspection reports. "Broken wire cages with sharp edges, that's a no-no. Collars that had cut into the skin—you are supposed to inspect them every day," Feldman says.

Still, Feldman says, "Were the animals suffering terribly? No. ... Fundamentally the guy wants to do right by the animals. He still hasn't come to the conclusion that right is going mean a big investment."

USDA has not filed a formal complaint against RCR, the other facility often used by researchers. But a USDA inspector cited RCR in 2017 for failing to disclose the existence of 1000 chinchillas and for using an unspecified "painful" and "unacceptable" method of euthanasia. In February and March, USDA inspectors reported RCR animals with eyes swollen or crusted shut, a cold and rigid dead chinchilla, and a filthy facility reeking of ammonia. The proprietor, Jan Ryerson, could not be located for comment.

AALAS will await the outcome of the USDA legal proceedings before taking any action to remove MCR from its Buyers Guide, Parker says. However, the University of Oklahoma reports that it dropped MCR as a supplier in 2018. "[MCR] animals were not of the quality we needed for our research," says Ronald Banks, the university's director of comparative medicine. ■



This chinchilla with untreated swelling under its chin was photographed at Moulton Chinchilla Ranch in 2017.

plier, Moulton Chinchilla Ranch (MCR) in Chatfield, Minnesota, has a 9-year record of violations and was to have faced an agency judge in April at a hearing; that hearing was postponed because of the pandemic.

MCR is the only chinchilla provider listed in the Buyers Guide produced by the American Association for Laboratory Animal Science (AALAS). "How is this business allowed to continue to operate given the issues that have been repeatedly identified?" asks Cathy Liss, president of the nonprofit

Welfare Act is written in such a way that it can be broadly interpreted for purposes of citation," Moulton told *Science*.

Long-tailed chinchillas (*Chinchilla lanigera*) are squirrel-size rodents native to the Andes. In 2019, U.S. biomedical researchers used 1250 of the animals, down by about one-third from the numbers used in 2014, according to USDA data curated by AWI. MCR and the other offending facility, Ryerson Chinchilla Ranch (RCR) in Plymouth, Ohio, are the suppliers most often cited in

FORCED INTO BATTLE

Bispecific antibodies unleash T cells against cancer
by physically tethering them to tumor cells

Amy Boland has gone through many ups and downs since she noticed lumps under her arms 12 years ago and learned she had cancer of the lymph system. For about 6 years, conventional chemotherapy helped shrink her lymphoma tumors, but they started to grow again. A succession of other cancer therapies, including a bone marrow transplant and a class of drugs called checkpoint inhibitors, either failed or only brought temporary relief. In one elaborate effort, physicians harvested her T cells, engineered those immune cells to kill her lymphoma, and infused them back into her body. The cancer vanished, but 2 years later bounced back. “Nothing was really working,” says oncologist Stephen Schuster of the University of Pennsylvania (UPenn).

So in October 2018, Boland joined a clinical trial testing another way to harness her immune system to kill the tumor cells. The idea of the trial, which Schuster co-leads, was to use a molecular rope known as a bispecific antibody to tether her natural T cells to the tumor cells so the immune warriors would attack. Like the engineered T cells she had received earlier, the experimental infusions sometimes made her sick enough to spend a couple of nights in the hospital. But the antibody rapidly sent her into remission. Today, more than 1 year after going off the Roche drug, mosunetuzumab, Boland, now 60, appears to be free of cancer and leads a normal life. “I’m feeling really good. I’m so grateful,” she says.

The ongoing trial Boland participated in made headlines in December 2019.

By **Jocelyn Kaiser**

At a meeting of the American Society of Hematology, Schuster reported that the antibody shrank fast-growing non-Hodgkin lymphoma tumors in 46 of 124 patients in whom other treatments had failed. For some of those people, like Boland, the failures included having their immune cells altered to attack the cancer. Those engineered cells, known as chimeric antigen receptor (CAR) T cells, have achieved remarkable results in some cancers. Yet at the same meeting, data from a small clinical trial suggested a bispecific antibody might work equally well on myeloma, another blood cancer. Bispecific antibodies for cancer are “superhot,” says Janice Reichert, executive director of the Antibody Society, who has tracked their development.

That’s the culmination of a slow boil. Scientists have worked on bispecific cancer drugs for decades and the first clinical success came 12 years ago. That result jump-started the field for a while, but other therapies, including CAR T cells, raced forward, in part because the cancer-fighting antibodies proved challenging to design and produce. But companies have now made those protein drugs safer and more potent. Variants are being tested in dozens of clinical trials, in the hope they can rival or surpass the engineered cells.

“If bispecific antibodies can do what CAR T cells can do, this would represent a big advance” and a “potentially fundamental change,” said Johns Hopkins Medicine hematologist Robert Brodsky at a press preview for the meeting where Schuster

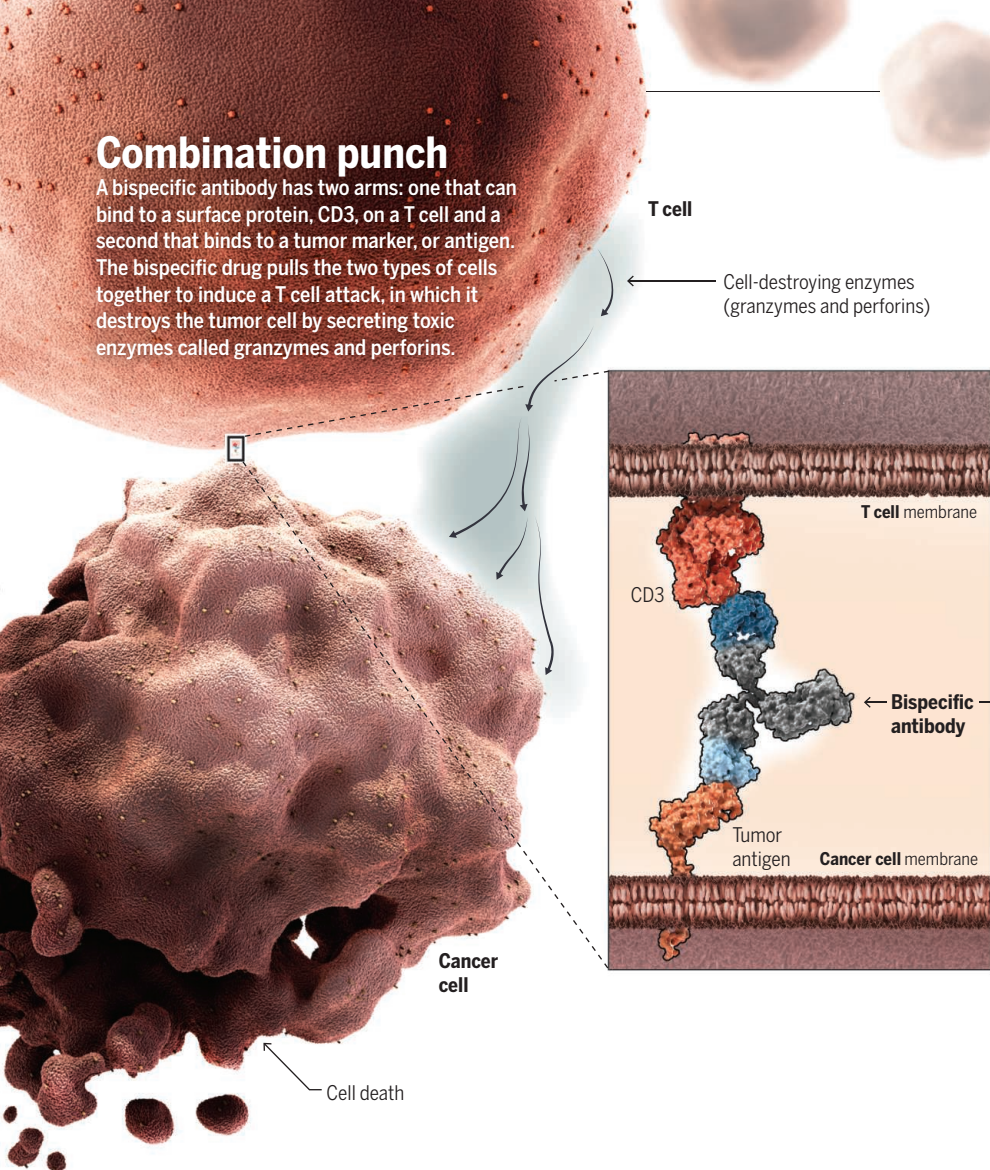
presented. A major advantage of bispecific antibodies is that they can be mass-produced in advance. CAR T cells, by contrast, must be prepared for each cancer patient. That process is costly and, for some very sick patients, takes too long.

Those new players in immunotherapy are no panacea yet. For some blood cancers, the bispecific antibodies aren’t giving patients the long-lasting remissions often seen with CAR T cells. As happened with CAR T cells, several patients have died in trials testing bispecific antibodies, possibly from overzealous immune responses sparked by the drugs. And bispecific antibodies may prove less effective against solid tumors, such as those of the colon and lungs, than against blood and lymph cancers—a drawback shared with CAR T cells. “There are a lot of open questions. But it’s also a field moving quickly, and a lot of really smart people are working on it,” says Paul Carter, an antibody researcher at Genentech, a Roche subsidiary.

ANTIBODIES HAVE A LONG HISTORY as a cancer treatment. The Y-shaped proteins are normally pathogen fighters, latching onto an antigen—a protein or a bit of one—on viruses, bacteria, or other microbes. The binding, which takes place at the tips of the Y, can directly disable and clear a pathogen or can signal the immune system to attack it. Cancer researchers first learned to exploit that natural system by making many copies of a specific antibody that latches onto an antigen unique to a particular cancer. This marks the cancer for destruction by com-

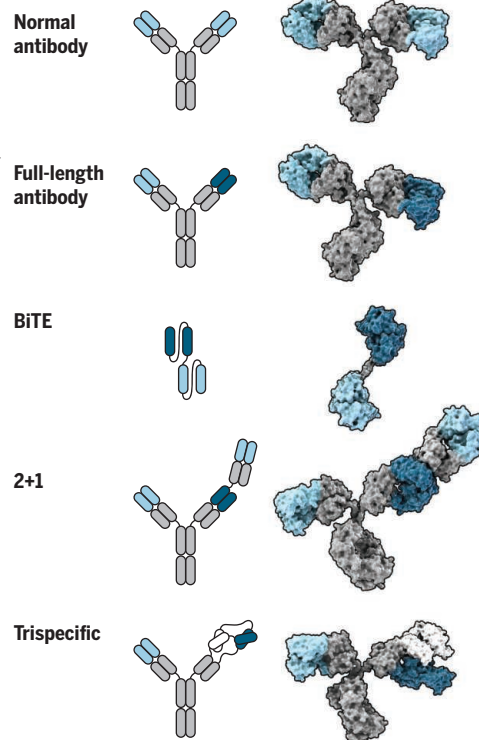
Combination punch

A bispecific antibody has two arms: one that can bind to a surface protein, CD3, on a T cell and a second that binds to a tumor marker, or antigen. The bispecific drug pulls the two types of cells together to induce a T cell attack, in which it destroys the tumor cell by secreting toxic enzymes called granzymes and perforins.



Design gallery

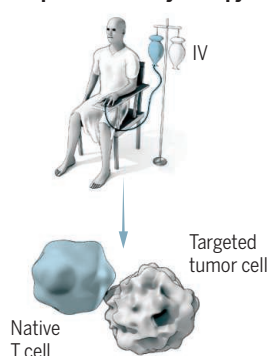
Arms of typical antibodies have identical tips (light blue) that can bind to a tumor. Bispecific antibodies that enlist T cells also have a binding site (dark blue) for the immune cells. Variants include a bispecific T cell engager (BiTE), which lacks a stem; a "2+1" antibody that has two sites that bind the same tumor antigen; and a trispecific, which binds two different surface proteins on a T cell.



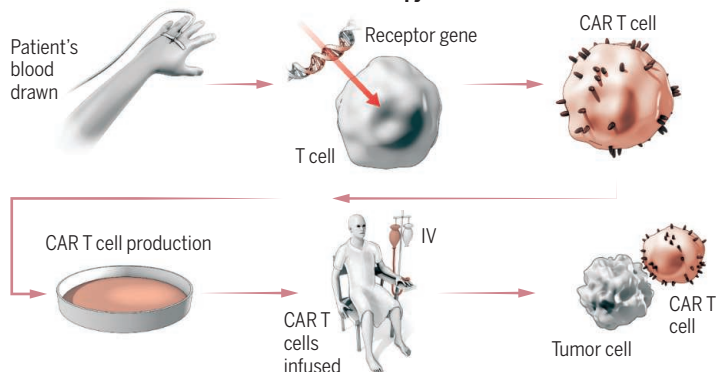
Dueling treatments

Bispecific antibodies are off-the-shelf drugs given intravenously. That's a simpler treatment than chimeric antigen receptor (CAR) T cells, made by harvesting T cells from a patient's blood, engineering them to have receptors for a tumor antigen, growing the cells, and infusing them back into the patient.

Bispecific antibody therapy



CAR T cell therapy



ponents of the immune system other than T cells. Some of the most effective, and best-selling, cancer drugs are such monoclonal antibodies, including the breast cancer drug trastuzumab, better known as Herceptin.

Newer anticancer strategies enlist T cells. Tumor cells can appear foreign enough for the body to sometimes "train" these cells to attack them. But CAR T cells, altered to carry a receptor that targets a cancer cell

antigen, can deliver a more powerful response. Checkpoint inhibitors, drugs that release the molecular brakes that can restrain T cells, can also boost the T cell attack (*Science*, 20 December 2013, p. 1432).

Bispecific antibodies offer a third way to harness T cells. In the mid-1980s, cancer researchers began to engineer antibodies that had two tips—one matched to a cancer cell antigen and the other to a T cell surface

protein called CD3. The idea was to directly link T cells to tumor cells, thereby skipping the need for T cells to learn to attack a cancer. "It's mimicking what naturally happens, but the advantage is that you can engage all T cells," not just those trained to attack the tumor, says Dirk Nagorsen, a vice president and cancer researcher at Amgen. In 1985, the field was galvanized by two reports in *Nature* that such a "bispecific" antibody

Amy Boland has fought lymphoma for more than a decade, but her cancer vanished after treatment with bispecific antibodies.



could destroy cancer cells in a dish; studies soon showed those antibodies could shrink tumors in mice.

The drugs were hard to make. Antibodies are modular, with two identical “heavy” chains, making up the stem and half of each arm of the Y, and two identical “light” chains, each of which completes one arm. Trying to assemble bispecific antibodies from those complex components, protein chemists got 10 versions of each molecule. That outcome meant laborious efforts to sift out the one researchers wanted.

And the excitement faded when tests of bispecific antibodies moved from lab animals and cells to cancer patients. In an early clinical trial, one antibody appeared to shrink lymphoma. But researchers had to stop treating patients before the study yielded definitive results because the antibody’s maker ran out of the drug. The

antibodies also sometimes triggered serious side effects, including liver damage and an immune overreaction in which white blood cells pump out signals called cytokines that can be toxic in large quantities. Such cytokine “storms” cause fevers and, in severe cases, organ damage. (CAR T cells can cause the same overreaction.)

Two immunologists at the Ludwig Maximilian University of Munich, Peter Kufer and Gert Riethmüller, pressed ahead anyway with an idea some colleagues thought would fail: a stripped-down bispecific antibody with two tips linked by a flexible peptide instead of the traditional stem. The simplified design made the antibody easier to manufacture, but because the stem was missing, the kidneys cleared it from the blood within 2 hours. In its first clinical test, for non-Hodgkin lymphoma, patients had to wear a pump to continually infuse

the antibody. Still, tiny doses of the drug shrank tumors in all seven lymphoma patients in the trial, run by Micromet, a German biotech Riethmüller co-founded. “We thought, ‘Oh my God, there’s something amazing happening here,’” says molecular biologist Patrick Baeuerle, then-chief scientific officer of Micromet, who dubbed the concept a bispecific T cell engager (trade-marked as BiTE).

The small trial, published in *Science* in 2008, sparked interest from companies and academics. “The whole field realized, ‘This is a big deal. We want in on this,’” says John Desjarlais, chief scientific officer of the biotech Xencor. At about the same time, CAR T cells began to show impressive results in some leukemia patients—which boosted interest in other, potentially simpler immunotherapies such as bispecific antibodies. Like CAR T cells, T cells stimulated by BiTEs

PHOTO: MARY BOLAND

release toxic molecules called granzymes and perforins that punch holes in tumor cells and cause them to self-destruct. “I see bispecifics [such as BiTEs] as an off-the-shelf CAR T cell,” says Elad Sharon, senior investigator with the National Cancer Institute’s Cancer Therapy Evaluation Program.

The first bispecific antibody for cancer was approved in Europe in 2009. It was meant to mop up the malignant cells that cause abdominal fluid to build up in some cancer patients—but it didn’t work that well, so the drug only stayed on the market a few years. The field regained momentum, however, after Amgen snapped up Micromet in 2012 and later showed that its BiTE drug, blinatumomab (Blinicyto), doubled the survival time of patients with advanced acute lymphocytic leukemia. Beginning in 2014, the Food and Drug Administration approved the drug to treat several adult and pediatric forms of the disease. Amgen is now testing BiTEs for other cancers, including myeloma and lung, prostate, and brain cancers.

OTHERS HAVE RUSHED to improve on BiTEs by using protein engineering tricks to create desired bispecific antibodies. Some firms have restored the antibodies’ stem, known as the Fc receptor, so that the protein stays in the blood longer—but with modifications that make it less toxic to the liver. Cancer patients such as Boland no longer have to wear a fanny pack containing a pump, but can now receive the drug as an intravenous drip every 3 weeks.

Industry scientists have also added a second copy of the tumor antigen-binding site to one tip of the antibody (see graphic, p. 931). Known as a “2+1” bispecific antibody, this particular design is meant to make the antibody more selective for cancer cells and less likely to target healthy cells carrying small amounts of the cancer antigen.

To reduce the risk of triggering a cytokine storm, researchers are also designing bispecifics that, instead of T cells, snag a different kind of immune cells called natural killers (NK). Several companies have started or are readying clinical trials of such antibodies, which bind to an NK cell surface protein called CD16. “NK cells are very potent tumor cell killers if activated, and the cytokine release is significantly reduced with this cell type,” says Dmitri Wiederschain, head of cancer immunology at one such company, Sanofi.

These variants and scores of other bispecifics are easier to synthesize now. “An-

tibody engineering has become so sophisticated, it’s possible to make these molecules quite efficiently,” says biochemist Christoph Rader of Scripps Research in Jupiter, Florida.

By Reichert’s count, more than 60 T cell-directing bispecific antibodies for cancer are in early- or later-stage clinical trials. One Amgen BiTE has shown hints of shrinking tumors in a few patients with advanced prostate cancer, the company reported last year.

Solid tumors are a challenging target for bispecifics in part because tumors often lack a unique antigen for the antibodies to grab. Many tumors are also surrounded by blood vessels, tissue, and immune cells that form a barrier T cells can’t easily penetrate. But findings from mouse studies suggest some bispecific antibodies can drive T cells into tumors, says Nai-Kong Cheung of Memorial Sloan Kettering Cancer Center. His lab has systematically tweaked design factors, such

the task by creating two bispecifics. One targets a tumor antigen and CD28 or another growth-signal receptor; the other binds to the tumor antigen plus CD3. “One of our hopes is that this costimulatory bispecific may help us unlock responses in solid tumors,” says Lowy, whose company reported in *Science Translational Medicine* in January that such a two-drug combination shrank ovarian tumors and slowed prostate tumor growth in mice.

Could those next-generation bispecifics eclipse CAR T cells for some cancers? UPenn’s Carl June, a CAR T cell pioneer, is skeptical. Many leukemia patients who get blinatumomab eventually relapse because cancer cells become resistant to the drug, he notes, so oncologists use it mainly as a “bridge” until a very sick patient can get a stem cell transplant or CAR T cells. June adds that bispecifics may not work in the many cancer patients whose T cells have become depleted

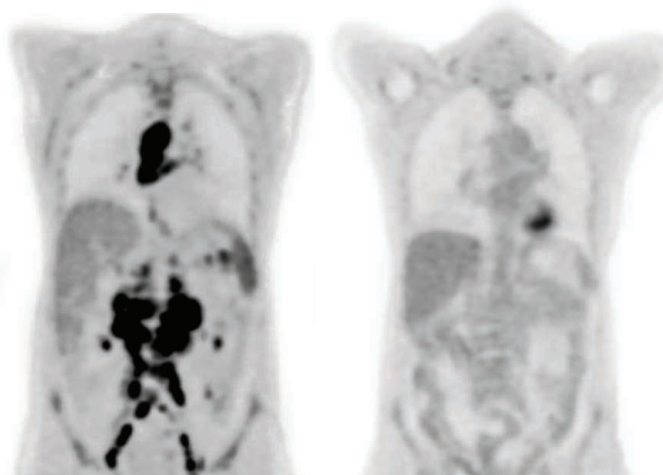
or “exhausted,” leaving too few to attack the cancer. CAR T cell treatment, by contrast, replenishes the immune ranks by growing the cells outside the body—something “not possible with bispecifics,” June says.

Schuster, who has a foot in both camps—he runs Boland’s study and has led CAR T trials with June—says bispecifics are still proving themselves. He points to an Amgen report last year that some lymphoma patients who responded well to blinatumomab were still alive after 7 years, suggesting they could stay in remission long term. “I am confident that within the next 2 to 5 years you’re going to see quantum leaps in our ability to target resistant tu-

ors, including solid tumors,” Schuster says.

He and other cancer researchers see CAR T cells, checkpoint inhibitors, and bispecific antibodies as interchangeable. “Why not do all of the above?” asks Schuster, who’s preparing a trial that will give patients CAR T cells and then a bispecific drug. “All these approaches to manipulate the cellular immune system to treat cancer are essentially different means to the same end.” Doubling up the treatments can be risky, however—last year, two patients died in such a trial after developing cytokine release syndrome.

Boland, who has seen all three of her children grow up and go off to college during her cancer fight, welcomes the progress. A bispecific drug is keeping her lymphoma in check for now, she notes. “I hope it’ll last, but if not, I feel confident that there’s always more treatments. You just don’t worry about it. Nobody knows what will happen.” ■



Amy Boland’s cancer scans showed a dramatic disappearance of her lymphoma tumors from the start of her bispecific antibody treatment (left) to 12 weeks later.

as how binding sites are arranged, to learn what optimizes the molecules’ potency.

And some companies hope to boost the attack on solid tumors with antibodies that bind not only to CD3, but also to another receptor on T cells known as a “second signal,” which stimulates the cells to grow. For years, says Regeneron Senior Vice President Israel Lowy, industry has been “afraid to touch” that protein, called CD28, because of a devastating mishap: An antibody designed to bind to it made six healthy volunteers critically ill from cytokine release syndrome in a 2006 U.K. clinical trial.

Findings from new studies, however, suggest it’s possible to exploit that cell growth trigger safely. Last year in *Nature Cancer*, a Sanofi team reported that a “trispecific” antibody with arms matched to CD28, CD3, and a cancer antigen wiped out myeloma tumors in mice. Other firms have split up

INSIGHTS

PERSPECTIVES

ECOLOGY

Exotic plants get a little help from their friends

Interactions with herbivores and microbes link exotic-plant success with carbon cycling

By **Carlos Urcelay**¹ and **Amy T. Austin**²

Terrestrial ecologists have identified multifaceted controls—climate, biogeography, disturbances, and their interactions—that shape how plant communities in natural ecosystems organize in space and time. Multiple documented interactions directly link plant diversity with other biotic guilds (herbivores, root symbionts, bacteria, and pathogens) and ecosystem processes [carbon (C) and nutrient cycling] (1). However, all appears to go awry when exotic (non-native) plant species invade and establish themselves without human intervention; such changes affect the functioning and diversity of natural ecosystems (2). On page 967 in this issue, Waller *et al.* (3) provide insight into pathways that explain the underlying relationship between

plant invasions and acceleration of a crucial ecosystem process: C turnover.

Changes in plant diversity, species dominance, or nutrient cycling as a result of exotic plant invasion can alter ecosystems so substantially that they may be virtually unrecognizable from their original state. So what makes a non-native species a successful invader? Can scientists predict which species will have large impacts on ecosystems? It might seem straightforward to identify characteristics of success for exotic invader species, including faster growth rate, thinner leaves, or root anatomy (4–6), often with consequences for accelerated C turnover (2). However, definitive conclusions as to what traits predict invasive success remain elusive; this is due in part to the interactions between above- and belowground components of an invaded

ecosystem (including plants with contrasting life-history strategies) seldom having been examined simultaneously. Using a comprehensive, manipulable, outdoor experimental system (mesocosm), Waller *et al.* revealed that as the proportion of exotic plants increases, interactive effects from herbivore and soil-microorganism communities alter C cycling and other processes. Thus, changes in C cycling are indirectly mediated by the plant invaders through changes in these biotic interactions.

In an extensive experimental effort, the authors used plants and soils from subalpine grassland in New Zealand to establish 160 experimental outdoor ecosystems (mesocosms) that contained 20 singular combinations of eight plant species that varied in their proportions of exotic and native woody plants (see the photo). The plant

PHOTO: WARWICK ALLEN



communities were grown in soils previously conditioned by them (“home”) and in soils conditioned by plant species not present in the community (“away”), a technique that is commonly used to evaluate the importance of plant-soil feedbacks (6). Last, herbivore interactions were evaluated by adding invertebrate herbivores (such as leafhoppers, aphids, moths, beetles, and slugs) to half of the mesocosms. With plant traits as covariates, linear mixed-effects models were used to test how the proportion of exotics, herbivores, and soil treatments affected biotic

interactions and ecosystem processes. In addition, structural equation modeling was used to explore how exotic-plant functional types (those with different growth forms such as grasses, shrubs, and trees) and various other plant traits influence C cycling directly and indirectly (through modifying herbivore and soil-microorganism biomass).

A noteworthy aspect of the new study is the use of different functional types of plants, including woody species with various symbiotic strategies. This is relevant because the success of exotic plant species and their effects on new environments might be strongly determined by their symbiotic interactions (7). Chief among them are mycorrhizal fungi and nitrogen-fixing bacteria that associate with plant roots, providing access to soil nutrients and stress alleviation (8). The insights observed by Waller *et*

Mini-ecosystems (aerial view, left; inside a single mesocosm, right) allow interactions among plants, herbivores, and soil biota to be manipulated.

al. likely derive not from C-cycle changes that occur with exotic-plant dominance but from the mechanistic underpinnings that reveal why these changes occur.

Whereas plant richness (number of species) decreased with an increasing proportion of exotics, likely through competitive exclusion of natives, total plant biomass remained constant. Key to understanding the impacts of exotic plants on biotic interactions and C cycling is the widely explored plant trait that expresses the thickness of the leaves [called specific leaf area (SLA)], which is linked to herbivore palatability and nutrient-conservation strategies (9). Community-weighted SLA increased, at the community scale, with a growing presence of exotic plants and appears to have mediated some of the exotic-plant effects on biotic interactions and C cycling. For example, the increasing proportion of exotic species was accompanied by an increase in herbivore biomass, which could be explained by the presence of exotic nitrogen (N)-fixing herbs of high nutrient content and higher SLA, both of which imply enhanced leaf palatability (10). Soil microbial guilds also were affected. The biomass of arbuscular mycorrhizal fungi (AMF) decreased with an increase in the proportion of nonmycorrhizal exotic plant species with high SLA. An increase in bacterial biomass was associated with symbiotic N-fixing exotic woody species.

There is an emerging understanding of how the composition of soil fungi and bacteria can affect soil respiration (carbon dioxide production from soil microorganisms and roots) (11). In the Waller *et al.* study, the increased proportion of exotics triggered an increase in total soil respiration but only in the presence of herbivores and microbial guilds in the away soil. This process was positively influenced by the presence of exotic N-fixing woody plants and high community-weighted SLA. In turn, basal plant respiration (that which occurs without plant roots) increased in the presence of away soil biota, which was attributed to the activity of saprophytic fungi (non-AMF). The combination of exotic-plant traits and changes in the relative abundance of soil microbial guilds boosted nutrient and C cycling. Why total respiration increased only with aboveground herbivory remains unknown but reveals the complex pathways that drive plant-soil interactions.

The authors also showed that the increasing proportion of exotic plants accelerated decomposition of a standard substrate of black tea leaves indirectly through the exotic

¹Instituto Multidisciplinario de Biología Vegetal, Consejo Nacional de Investigaciones Científicas y Técnicas—Universidad Nacional de Córdoba (CONICET-UNC), UNC, Córdoba, Argentina. ²Instituto de Investigaciones Fisiológicas y Ecológicas Vinculadas a la Agricultura (IFEVA), CONICET, Universidad de Buenos Aires, Capital Federal, Argentina. E-mail: curcelay@imbiv.unc.edu.ar

plants' effects on soil fungi and herbivory. The increases in non-AMF biomass could be explained by the increasing presence of exotic N-fixing woody plants, together with reductions in nonsaprophytic AMF biomass by exotic non-N-fixing plants. This evidence provides support for the importance of life-history strategies in structuring below-ground fungal communities (12) and their interactions with herbivores. However, it is difficult to draw conclusions about direct links between ecosystem changes and exotic-plant traits because litter decomposition was evaluated with tea leaves from species not included in the experiments. The chemical and morphological characteristics of plant-litter identity (13) could contribute to the complexity of the patterns observed in the new study. Thus, the complex linkages among plant invasions, soil biota, and litter decomposition deserve further exploration.

The new work by Waller *et al.* highlights the fundamental role of biotic interactions in C-cycling processes in an experimental setting. It remains unclear how these results will scale up to field communities with more complex biological networks. Future research will unearth whether these interactive pathways are maintained or accentuated with adult woody species in invaded natural ecosystems or in other invasive scenarios, such as those lacking N-fixing invaders or those dominated by other mycorrhizal types. What seems evident from this study is that linkages between plant invasions and ecosystem processes are of pivotal importance for understanding the interplay of how plants interact with other organisms (14). Researchers must delve further into this complexity to better predict how these links will respond to global changes. ■

REFERENCES AND NOTES

1. W. W. Weisser *et al.*, *Basic Appl. Ecol.* **23**, 1 (2017).
2. J. G. Ehrenfeld, *Annu. Rev. Ecol. Syst.* **41**, 59 (2010).
3. L. P. Waller *et al.*, *Science* **368**, 967 (2020).
4. P. A. Tecco, S. Díaz, M. Cabido, C. Urcelay, *J. Ecol.* **98**, 17 (2010).
5. I. Jo, J. D. Fridley, D. Frank, *Biol. Invasions* **17**, 1545 (2015).
6. R. A. Wilschut *et al.*, *Nat. Commun.* **10**, 1564 (2019).
7. M. P. Thakur, W. H. van der Putten, M. M. P. Cobben, M. van Kleunen, S. Geisen, *Nat. Rev. Microbiol.* **17**, 621 (2019).
8. F. M. Martin, S. Uroz, D. G. Barker, *Science* **356**, eaad4501 (2017).
9. S. Díaz *et al.*, *J. Veg. Sci.* **15**, 295 (2004).
10. M. Schädler, G. Jung, H. Auge, R. Brandl, *Oikos* **103**, 121 (2003).
11. J. Whitaker *et al.*, *J. Ecol.* **102**, 1058 (2014).
12. D. Francioli *et al.*, *Plant Soil* (2020). 10.1007/s11104-020-04454-y
13. L. Vivanco, A. T. Austin, *J. Ecol.* **96**, 727 (2008).
14. A. T. Austin, C. L. Ballaré, *New Phytol.* **204**, 257 (2014).

ACKNOWLEDGMENTS

This work was supported by CONICET-UNC, the Secretariat of Science and Technology, and Ministerio de Ciencia de la Provincia de Córdoba (C.U.) and the New Phytologist Trust (A.T.A.)

10.1126/science.abc3587

SYNTHETIC BIOLOGY

Remote activation of cellular signaling

Electric fields stimulate genetically modified pancreatic β cells to secrete insulin in mice

By **Matthew I. Brier**¹
and **Jonathan S. Dordick**^{1,2}

Remote modulation of cell signaling has biological ramifications across the cellular, tissue, and organismal levels, with applications as diverse as medicine and biomanufacturing.

External approaches to regulating biological processes invariably require genetic manipulation. Platforms with orthogonal means of remotely controlling cell function include optogenetics (1, 2), mechanogenetics (3, 4), and magnetogenetics (5–7). An array of techniques make use of various pathways to control cell signaling and enable in-depth studies of cell processes and the development of therapeutic tools. An external stimulus can control one or more cellular pathways leading to, for example, control of ion channels, temporal production of proteins, or neural circuit modulation. On page 993 of this issue, Krawczyk *et al.* (8) advance the use of electrogenetics, a method that uses electric fields to control cell function, through the development of a wearable device that releases insulin from bioengineered cells and controls blood glucose concentrations in vivo.

Electrogenetics uses cells that have been engineered to respond to electrical stimulation through the expression of voltage-dependent receptors, such as voltage-gated ion channels. Krawczyk *et al.* exploited a voltage-gated calcium channel (Ca_v1.2) coupled to an inwardly rectifying potassium channel (K_{ir}2.1) to complete a voltage-gated circuit and provide a high degree of control over gene expression in human embryonic kidney cells (see the figure, top left). Using voltage-controlled square pulses, the authors demonstrated electrostimulation-driven expression of a simple marker protein, secreted embryonic alkaline phosphatase, with the amount expressed controlled by the applied voltage, pulse length, and duration of stimulation.

Having demonstrated model protein production at the transcriptional level, Krawczyk *et al.* then engineered a similar set of genetic constructs into a monoclonal pancreatic β cell line (Electro β cells) with vesicular insulin secretion functionality decoupled from glucose sensitivity. This modification enabled insulin to accumulate in the Electro β cell vesicles and to be released upon electrostimulation, rather than in response to changes in glucose concentrations. Peak secretion occurred within 10 min after starting stimulation. The authors extended these in vitro results to a diabetic mouse model by encasing Electro β cells in an electrostimulation device that could be subcutaneously implanted. Implants facilitated reuse of the Electro β cells for several weeks. Wireless activation could induce electrostimulation and secretion of stored insulin to counteract perturbations in blood glucose concentrations within minutes.

This electrogenetics approach shows capabilities similar to those of optogenetics, perhaps the most well-studied means to control cell function remotely, in which precise wavelengths of light stimulate wavelength-specific light-sensitive receptors and ion channels (see the figure, top right) (1, 2). The discovery of bacterial channelrhodopsin (ChR) as a calcium ion-conducting photoreceptor (9) led to myriad applications, including control of neuronal action potentials with millisecond-time scale spatiotemporal control (2). Mutations made to ChR family members have introduced new functionalities, including on-off switching and Cl[−] conductance (10). However, the poor tissue penetration depth of light requires fiber-optic implants, although recent studies with near-infrared light to activate lanthanide-doped nanoparticles to emit visible light may help to overcome accessibility restrictions (11).

Alternatively, mechanogenetics involves physical stimuli that have deep tissue penetration depths, such as focused ultrasound (FUS), to actuate mechanosensitive receptors. As an example, two-pore domain potassium ion channels genetically introduced into rat hippocampal neurons in vitro were activated by FUS to drive membrane depolarization and action potentials (3). In a related approach, microbubbles decorated

¹Department of Chemical and Biological Engineering and Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY 12180, USA.

²Departments of Biomedical Engineering and Biological Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, USA. Email: dordick@rpi.edu

with streptavidin and targeted to biotinylated Piezo1 mechanosensitive channels were used to transduce 150-MHz ultrasound and gate Ca^{2+} flux into chimeric antigen receptor T cells, thereby controlling Ca^{2+} -dependent gene expression as a potential cancer therapy (see the figure, bottom left) (4).

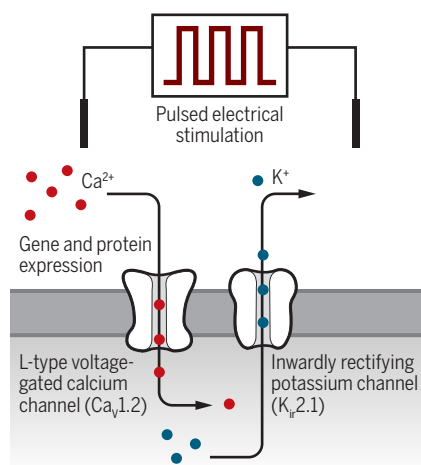
Another remote physical technique with similar spatiotemporal control and tissue penetration is magnetogenetics, which uses time-variant magnetic fields to activate magnetic nanoparticles targeted to cell receptors. Early studies in magnetogenetics used external magnetic iron oxide nanoparticles localized to target receptors, such as the transient receptor potential (TRP) vanilloid 1 (TRPV1) cation channel, that were exposed to magnetic stimulation to control Ca^{2+} entry into cells (5). More recently, several variations of this platform have been developed using genetically encoded ferritin, an endogenous cage protein responsible for maintaining iron homeostasis by sequestering iron as a magnetic nanoparticle, engineered to bind to TRP channels (see the figure, bottom right) (6, 7). Exposure of cells expressing the ferritin-conjugated TRPV1 platform to 465-kHz alternating magnetic field (AMF) stimulation allowed control of Ca^{2+} -dependent insulin production in vitro (6) and of blood glucose concentrations and feeding behavior in mice (12).

Despite promising results, magnetogenetics has been the subject of vigorous debate over the underlying mechanism and the veracity of the results to date. Although localized heating or mechanical torque in the vicinity of the exogenously added nanoparticles were viewed as logical means to gate TRP channels, arguments targeting deficiencies in ferritin's magnetic properties suggested that such mechanisms could not explain the observed phenomenon (13). Nonetheless, alternative mechanisms have been proposed, including magnetocaloric heating resulting from changes in the magnetic entropy of the iron biomineralized core within ferritin that could conceivably gate TRP channels (14). Chemical mechanisms have also been advanced that attribute gating of TRP channels to iron release from ferritin in the presence of AMFs; free iron results in localized formation of reactive oxygen species and lipid oxidation of the cell membrane in the channel's vicinity, both of which gate TRP channels (15).

We can now add electrogenetics to the mix of technologies for rapid and remote spatiotemporal control over complex biological functions. Electrogenetics represents the next tool in an expanding toolbox for engineering remote solutions for human therapeutics. Despite their inherent limitation, each of these platforms has shown applicability, and the potential for combinations of these orthogonal platforms to exert fine-tuned con-

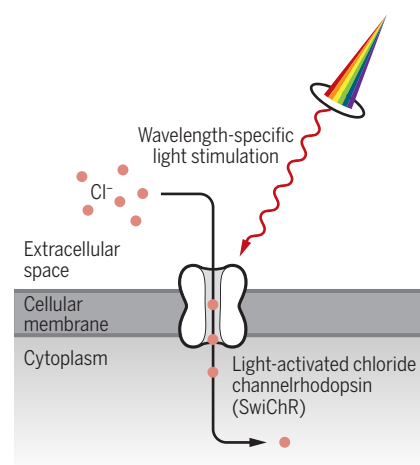
Remote control of cells

Krawczyk *et al.* report an electrogenetic approach to stimulate cells. This method joins three others for actuating genetically modified cell membrane ion channels to initiate signal transduction and control cell function.



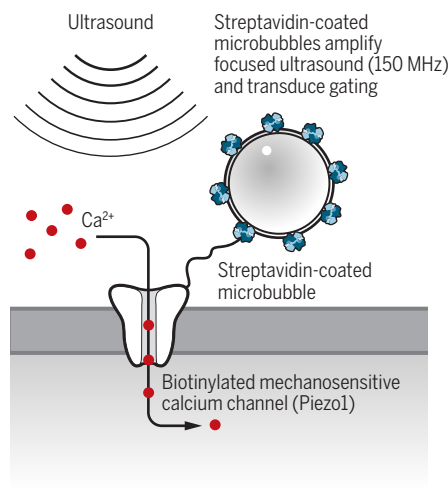
Electrogenetics

Krawczyk *et al.* use pulsed electrical stimulation to control a voltage-gated circuit. The change in Ca^{2+} concentration leads to expression of a simple marker protein, secreted embryonic alkaline phosphatase, and vesicular insulin secretion (8).



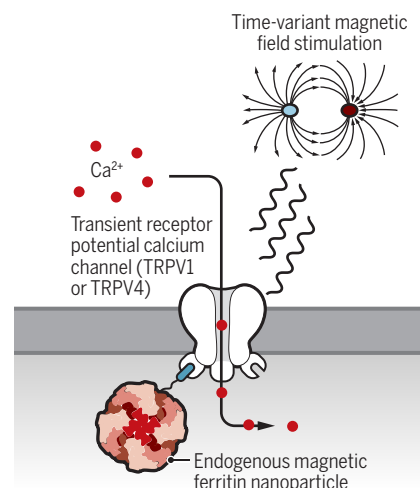
Optogenetics

Precise wavelengths of light stimulate light-sensitive receptors and ion channels, in this case SwiChR (Cl^- -conducting step-waveform inhibitory channelrhodopsin) (10).



Mechanogenetics

Streptavidin-coated microbubbles targeted to biotinylated Piezo1 mechanosensitive channels transduce focused ultrasound to gate Ca^{2+} flux into chimeric antigen receptor T cells (4).



Magnetogenetics

Exposing cells expressing ferritin-conjugated TRPV1 to an alternating magnetic field controls Ca^{2+} -dependent insulin production in vitro (6).

trol in complex systems is within our grasp. Research into synergistic use of these approaches could be the next step. For example, self-contained implants could act as multimodal monitoring and treatment devices for targeted diseases that have proven elusive to the current toolbox of modern medicine. ■

REFERENCES AND NOTES

1. F. Zhang *et al.*, *Cell* **147**, 1446 (2011).
2. E. S. Boyden *et al.*, *Nat. Neurosci.* **8**, 1263 (2005).
3. J. Kubanek *et al.*, *Sci. Rep.* **6**, 24170 (2016).
4. Y. Pan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 992 (2018).
5. H. Huang *et al.*, *Nat. Nanotechnol.* **5**, 602 (2010).

6. S. A. Stanley *et al.*, *Nat. Med.* **21**, 92 (2015).
7. A. A. Mosabbir, K. Truong, *ACS Synth. Biol.* **7**, 718 (2018).
8. K. Krawczyk *et al.*, *Science* **368**, 993 (2020).
9. G. Nagele *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13940 (2003).
10. A. Berndt *et al.*, *Science* **344**, 420 (2014).
11. S. Chen *et al.*, *Science* **359**, 679 (2018).
12. S. A. Stanley *et al.*, *Nature* **531**, 647 (2016).
13. M. Meister, *eLife* **5**, e17210 (2016).
14. G. Duret *et al.*, *Biophys. J.* **114**, 669a (2018).
15. M. Hernández-Morales *et al.*, *Cell Rep.* **30**, 3250 (2020).

ACKNOWLEDGMENTS

Supported by NIH grant R01 NS097184.

10.1126/science.abb9122

CANCER

Probing the tumor micro(b)environment

Bacteria are widespread in tumors, are found within cells, and differ by cancer type

By **Chloe E. Atreya**¹ and **Peter J. Turnbaugh**^{2,3}

Bacteria have been implicated in the initiation and progression of cancers originating on mucosal surfaces that either harbor a diverse microbial community (microbiota) or are routinely exposed to microbes from the environment (1–3). Far less is known about the potential for bacteria to influence tumors in body sites that are typically considered sterile. One hypothesis is that the abundant and diverse microbiotas found on mucosal surfaces may exert “remote control” by releasing small molecules into circulation (4, 5). An alternative, non-conflicting hypothesis is that the tumor microenvironment harbors microbes that exert local effects. This hypothesis is supported by the detection of bacteria in a growing number of tumor types (6, 7), although the reliability of distinguishing low-abundance bacteria from contamination has been questioned (8). On page 973 of this issue, Nejman *et al.* (9) present the most rigorous and comprehensive survey of bacteria in human tumor samples to date.

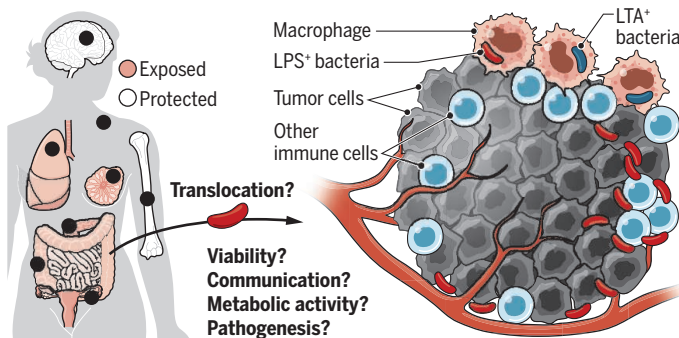
Nejman *et al.* use a new five-region 16S ribosomal RNA gene sequencing method, microscopy, and cell culture to characterize tumor-residing bacteria at known and previously uncharacterized sites. They report that most cancers harbor bacteria, albeit at low diversity except in breast cancer. Surprisingly, these bacteria appear to be intracellular within both cancer and immune cells. Moreover, they report associations between specific bacteria and tumor type and subtype, smoking status, and immunotherapy response.

These results raise multiple important questions for future study (see the figure). For example, is the level of diversity and

physiological status of these bacterial cells sufficient to constitute a “microbiota”? Although the defining characteristics of microbiotas remain in flux, two general themes are extensive microbe-microbe and host-microbe interactions, often over long time scales. Are the tumor-residing bacteria found within human cells able to communicate with each other? Prior work (6) suggests that bacteria found in tumors can metabolize drugs; however, the overall viability and metabolic activity of tumor-residing bacteria are unclear.

The hidden microbiota inside cancer

Intratumoral bacteria have been detected in both mucosal (exposed) body sites and protected sites. Nejman *et al.* examined bacterial occurrence in multiple primary tumor sites (black dots) and found that tumor cells and immune cells may harbor lipopolysaccharide-expressing (LPS⁺) bacteria, whereas macrophages contain at least remnants of LPS⁺ and lipoteichoic acid-expressing (LTA⁺) bacteria. These findings raise numerous questions that require further study.



The stability of bacterial diversity in tumors remains to be determined. Are tumor-residing bacteria seeded early on in tumorigenesis, or does the tumor alter the microenvironment such that bacteria can continually invade? In mouse models of pancreatic cancer (7), the gut microbiota appears to determine which bacteria are found in tumors, suggesting that there is a potential for bacteria to migrate into tumors at later stages. Longitudinal studies in patients with paired analyses of microbial diversity in tumors and mucosal surfaces are an important next step.

Why are bacteria found in tumors? One possibility raised by Nejman *et al.* is that there are always low amounts of bacteria in human tissue, which is supported by analyses of matched normal adjacent tis-

sue from the mammary glands of patients and even healthy controls (9). In precancerous conditions, the detection of enterotoxigenic bacteria may portend a generalized procarcinogenic inflammatory state (10). After tumorigenesis, disruptions to physical and molecular barriers, together with relative immunosuppression, may increase the potential for bacterial translocation to sites that are normally sterile. This “leakiness” of the tumor microenvironment has been extensively described in the context of vascular permeability; however, the degree to which leakiness enables bacterial invasiveness remains unclear. Additionally, the observation that tumor-associated bacteria are intracellular raises the possibility that the bacteria do not actually move freely into tumors or adjacent tissues—they may be transported there, intact or in fragments, through the migration of immune or cancerous cells.

A key barrier to progress is the lack of representative models for studying tumor-residing bacteria or other low-biomass microbial communities (8). Studies of colorectal cancer demonstrate the persistence of viable *Fusobacterium* over successive passages of human tumors in immunodeficient mice (11). However, laboratory mice harbor microbial communities

(12) and immune profiles that are distinct from those found in humans. Development of “triple-humanized” models wherein cancer, immune, and microbial cells are transplanted from patients into germ-free mice may be necessary.

These models could help to untangle the relationships between tumor-residing bacteria and treatment response. By contrast, there is abundant evidence that gut bacteria modulate the immunotherapy responsiveness of cancers, even at distant sites (4). Immune cells in the tumor microenvironment also play major and actionable roles, but does this extend to tumor-associated microbes? For a given cancer type, it will be important to determine the contribution of microbial composition relative to other tumor cell intrinsic and

¹Department of Medicine, Division of Hematology and Oncology, and Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, CA, USA.

²Department of Microbiology and Immunology, G.W. Hooper Research Foundation, University of California, San Francisco, CA, USA. ³Chan Zuckerberg Biohub, San Francisco, CA, USA. Email: peter.turnbaugh@ucsf.edu

extrinsic factors that drive malignancy.

The conceptual shift toward studying bacteria within tumors provides challenges and opportunities for translational research. Unlike the microbiotas found on mucosal surfaces, tumor-residing bacteria are not readily manipulatable. Current options for microbiota modulation rely on dietary, pharmaceutical, and microbiological perturbations (13). It remains unknown if tumor-residing bacteria depend at all on dietary substrates or if they subsist entirely on host-derived nutrients. Targeting intracellular tumor-residing bacteria also poses drug-delivery challenges; it may be possible to co-opt antibody-drug conjugates or other methods to specifically target bacteria. Although there is a long history of delivering viable bacteria to tumors (14), the risks and benefits of this approach need to be carefully considered.

Nejman *et al.* emphasize that diverse bacteria are found on and in the human body. Continued investigation may benefit from the rich history of research on intracellular bacteria in insects and plants. Intracellular bacterial pathogens harbor elaborate machinery to manipulate host cellular pathways (15); it will be interesting to see if tumor-residing bacteria encode similar effectors that enable their survival and dissemination. Achieving a comprehensive understanding of the tumor microenvironment is a daunting yet critical step toward an organism-wide mechanistic model of cancer progression and, if successful, may unlock the next wave of precision cancer diagnostics and therapeutics. ■

REFERENCES AND NOTES

1. D. Börnigen *et al.*, *Sci. Rep.* **7**, 17686 (2017).
2. A. D. Kostic *et al.*, *Genome Res.* **22**, 292 (2012).
3. K. L. Greathouse *et al.*, *Genome Biol.* **19**, 123 (2018).
4. V. Gopalakrishnan *et al.*, *Science* **359**, 97 (2018).
5. E. N. Bess *et al.*, *Nat. Microbiol.* **5**, 56 (2020).
6. L. T. Geller *et al.*, *Science* **357**, 1156 (2017).
7. E. Riquelme *et al.*, *Cell* **178**, 795 (2019).
8. M. C. de Goffau *et al.*, *Nature* **572**, 329 (2019).
9. D. Nejman *et al.*, *Science* **368**, 973 (2020).
10. C. M. Dejea *et al.*, *Science* **359**, 592 (2018).
11. S. Bullman *et al.*, *Science* **358**, 1443 (2017).
12. J. Walter, A. M. Armet, B. B. Finlay, F. Shanahan, *Cell* **180**, 221 (2020).
13. K. N. Lam, M. Alexander, P. J. Turnbaugh, *Cell Host Microbe* **26**, 22 (2019).
14. M. Sedighi *et al.*, *Cancer Med.* **8**, 3167 (2019).
15. S. M. Moss *et al.*, *Cell Host Microbe* **25**, 454 (2019).

ACKNOWLEDGMENTS

Support is provided by NIH (R21CA227232) and the Damon Runyon Cancer Research Foundation (DRR-42-16). C.E.A. is on the scientific advisory board for Pionyr Immunotherapeutics and has research funding from Bristol Myers Squibb, Guardant Health, Merck, and Novartis. P.J.T. is on the scientific advisory boards for Kaleido, Pendulum, Seres, and SNIPR Biome.

GEOLOGY

Tracking the rapid pace of a retreating ice sheet

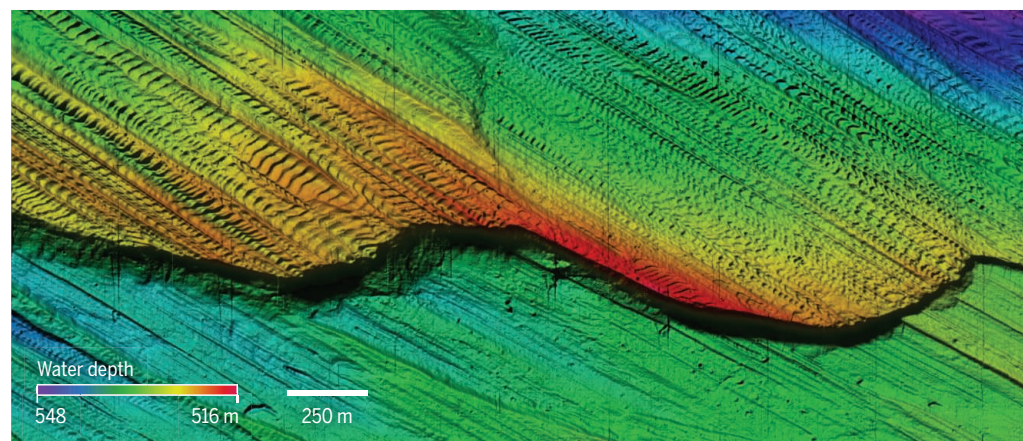
Seafloor mapping shows that Antarctic ice sheets retreated faster during the last deglaciation than today

By Martin Jakobsson^{1,2}

Glaciers and ice sheets that extended from land into the ocean left traces behind on the seafloor called submarine glacial landforms. If mapped in sufficient detail and interpreted correctly, they can provide comprehensive information into past behaviors of glaciers and ice sheets. On page 1020 of this issue, Dowdeswell *et al.* (1) describe the mapping of glacial landforms in the seafloor created by a rapidly retreating ice sheet on the eastern Antarctic Peninsula. The high-resolution data suggest that the retreat rate was paced by ocean tides and at least an order of magnitude faster than modern rates observed in other sensitive areas, such as West Antarctica where the ice sheet drains into the ocean at several locations (2). The retreat on the eastern Antarctic Peninsula took place more than 10,000 years ago, pointing out the challenges in predicting the sea-level rise contribution from retreating glaciers and ice sheets in a warming climate.

Glacial landforms have long been used to reconstruct ice-sheet extent and, specifi-

cally, its retreat pace and dynamics. At a meeting in Stockholm 1889 (3), the Swedish geologist Gerard De Geer presented observations of moraines, a general term to describe glacial landforms that consist of a mixture of debris (mainly sediments and rock fragments) deposited and sometimes molded by glacier ice. The moraines were observed in Sweden northwest of Stockholm. They generally took the form of ridges, between 1 and 5 m high and a few to slightly more than 10 m wide, gently winding through the landscape for several kilometers. There were several parallel lines of the moraines mapped in the landscape 200 to 300 m apart. De Geer compared the distances between the parallel moraines with the notion prevailing at the time that Swiss glaciers could retreat up to 70 m during 1 year. He put forward a hypothesis that the moraines were deposited during winter along the margin of the Scandinavian Ice Sheet when it made a seasonal halt during its retreat over the landscape. The distance between the moraines of 200 to 300 m represented therefore a yearly retreat rate of the ice sheet. About 11,000 years ago, when the Scandinavian Ice Sheet's margin was located in the area northwest of Stockholm, land was depressed below the contemporary water level of the Baltic Sea (4). This meant that the moraines described by



Seafloor imagery of the shape and depth of a grounding-zone wedge complex (where ice transitions from a grounded ice sheet to a floating ice shelf) was derived from an autonomous underwater vehicle-deployed multibeam echo-sounder as it surveyed part of Larsen Inlet, Antarctica. Grid cell-size is 1 m.

10.1126/science.abc1464

De Geer were formed at the ice margin under water. It would, however, take nearly a century after De Geer's presentation before geophysical seafloor mapping reached the technical capacity to be able to survey features with reliefs of only a meter or so. These "De Geer moraines" are similar in some aspects to the submarine glacial landforms from Antarctica mapped, described, and interpreted by Dowdeswell *et al.*

Dowdeswell *et al.* describe the appearance of the glacial landforms as if "ladders with numerous rungs" had been left embedded in the seafloor (see the figure). Similar features have been mapped and described previously, although often at lower resolution, around Antarctica (5–7) and in the Arctic (8). However, subtle differences with respect to appearances and settings led to different interpretations of the formation mechanisms. What is common to all previous studies and that of Dowdeswell *et al.* is that the proposed formation mechanisms involve ocean tides as a "pacemaker." The glacier ice, either of an iceberg or the margin of an outlet glacier draining an ice sheet into the ocean, can be slightly lifted from the seafloor during high tide. When settling again during low tide, a small ridge is formed on the seafloor by squeezing of sediments along the bottom edge of the iceberg or the outlet glaciers' margin. The submarine glacial landforms mapped in previous studies, referred to as "corrugation ridges" or "wash-

board patterns," were interpreted to have been formed in the wake of large grounded icebergs (7), or at the trailing end of drifting armadas of mega-sized icebergs calved during an ice shelf break-up (5), or underneath a thick ice shelf that floated close above the seafloor but bearing a protruding ice keel that could ground periodically with the tidal movement (6). All of these proposed formations of small, extremely regular ridge-like features in seafloor sediments take place at the tail end of the ice or by a bulging ice keel, moving away from the newly formed ridge without running over and destroying them. The ladders and rungs, by contrast, are formed at the front of the grounded ice margin, implying that when the ice moves forward, the features will be destroyed. Glacier ice moves by gravity from higher elevation on land toward the ocean, implying that the ice margin's grounding line must back-step for every formed rung. This can only happen if mass is lost at a regular pace. The precise ice dynamics of this regular mass loss that is paced by tide is not resolved and presents a challenge for the ice-modeling community. Although the mechanical details are yet to be characterized, Dowdeswell *et al.* have used the idea that the ocean tide-paces the formation of the rungs to convert the distances between the rungs to a retreat rate of the ice sheet.

Perhaps most importantly, Dowdeswell *et al.* demonstrate the immense value of

high-resolution seafloor mapping in unraveling the complex history of glacial dynamics. Only fractions of the seafloor in the hard-to-access ice-covered polar regions are mapped, and much is left to discover and learn. Scientists working on submarine glacial landforms have experienced a breakthrough in seafloor mapping capabilities similar to what their colleagues working in the terrestrial realm achieved when high-resolution terrain models of the land surface became available through airborne and satellite technology. Indeed, Dowdeswell *et al.* used the latest geophysical mapping technology mounted on an unmanned mini-submarine to survey the seafloor. Nonetheless, the pace of mapping the oceans, specifically in remote areas, is orders of magnitude slower because of the logistical difficulties and the fact that the ocean is in the way. ■

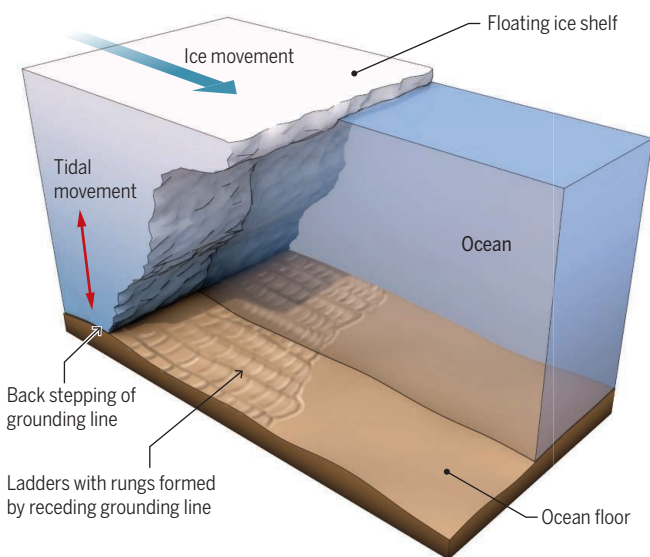
REFERENCES AND NOTES

1. J. A. Dowdeswell *et al.*, *Science* **368**, 1020 (2020).
2. E. Rignot, J. Mouginot, M. Morlighem, H. Seroussi, B. Scheuchl, *Geophys. Res. Lett.* **41**, 3502 (2014).
3. E. Erdmann, *Geol. Foren. Stockh. Forh.* **11**, 73 (1889).
4. S. Björck, *Quat. Int.* **27**, 19 (1995).
5. M. Jakobsson *et al.*, *Geology* **39**, 691 (2011).
6. A. G. C. Graham *et al.*, *J. Geophys. Res. Earth Surf.* **118**, 1356 (2013).
7. R. Lien, A. Solheim, A. Elverhoi, K. Rokoengen, *Polar Res.* **7**, 43 (1989).
8. K. Andreassen, M. C. M. Winsborrow, L. R. Bjarnadóttir, D. C. Rutherford, *Quat. Sci. Rev.* **92**, 246 (2014).

10.1126/science.abc3583

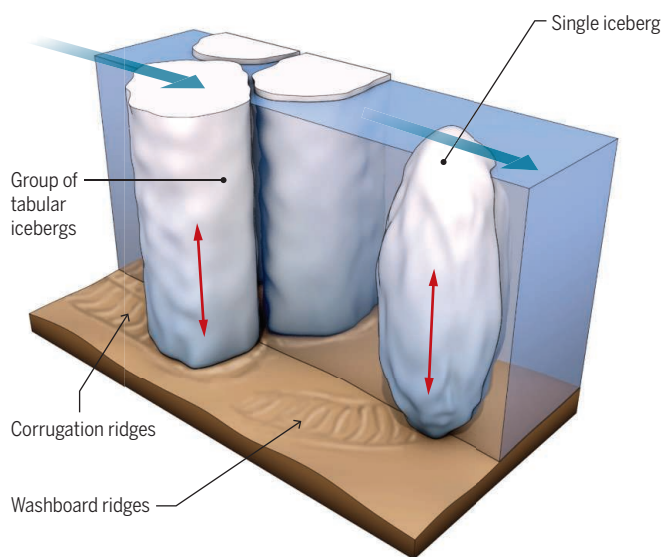
Unraveling glacial dynamics

Seafloor patterns can reveal grounded ice margin (left), iceberg (right), and under ice (not shown) scenarios.



Ladders with rungs

Ridges form when ice settles during low tide. The ice margin is lifted during the next high tide, loses mass from calving and melting, and settles 20 to 25 m further inland to form a ridge during the next low tide. If the ice sheet has a protruding ice keel that extends into a cavity below the grounded ice, it may form a different pattern (corrugation ridges).



Corrugation and washboard ridges

These patterns are created by icebergs moving individually or in a group. When the ice is moved up and down by the ocean tide, sediment ridges are squeezed out at their trailing edges.

Catching up to nature's ribosomes

Flow chemistry enables total chemical synthesis of proteins without solution-phase ligation

By **Caroline Proulx**

Chemical protein synthesis allows the systematic incorporation of one or more unnatural amino acids at precise locations along a protein backbone that enable structure-function relationship studies (1, 2). However, solid-phase peptide synthesis (SPPS) (3) is usually limited to peptides of 50 residues, and as a result, chemical protein synthesis requires native chemical ligation (NCL) (4) reactions for the chemoselective coupling of unprotected peptide fragments in solution. This combined multistep approach typically requires the synthesis and sequential ligation of at least three peptide fragments to access proteins, a process that remains relatively laborious. On page 980 of this issue, Hartrampf *et al.* (5) chemically synthesized single-domain proteins in hours. They circumvent the use of solution-phase chemistry by carefully optimizing iterative amino acid couplings on a solid support using a flow-chemistry setup (see the figure).

The introduction of SPPS in the 1960s (3) revolutionized peptide sciences. The attachment of the carboxyl-terminal amino acid to an insoluble polymeric support allowed each reaction to be done with excess reagents that could be easily washed off through simple filtration, a technique that earned Bruce Merrifield the Nobel Prize in Chemistry in 1984. Because there are no purifications steps in between each amino acid coupling and deprotection reactions performed on a solid support, near-quantitative yields are required for each monomer addition cycle to enable the synthesis and purification of proteins. Common by-products that have plagued peptide syntheses have included those arising from deletions, truncations, aspartimide formation, and amino acid epimerizations. Each unwanted reaction yields material that is often difficult or impossible to separate from the desired product with high-performance liquid chromatography purification methods.

Although much progress has been made to address these challenges (6), peptide sequence lengths that can be routinely accessed with traditional SPPS approaches fall well below the lengths of entire protein

sequences. This limitation has required the development and use of postsynthetic ligation methods. However, a wide variety of unnatural amino acids can be incorporated using this approach. Alternatively, incorporation of unnatural amino acids into proteins can be done through genetic code expansion in biological systems instead (7), albeit with differences in substrate scope.

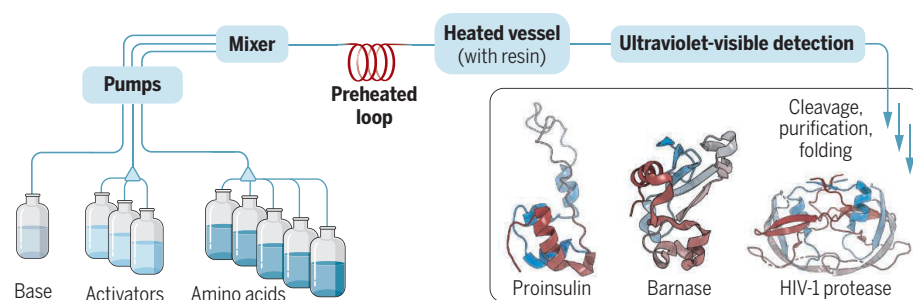
In a previous study from this team, Mijalis *et al.* reported automated fast-flow peptide synthesis (AFPS) procedures that substantially accelerated the process and improves the synthesis of difficult peptide sequences (8). Notably, their AFPS system has shortened the amino acid addition

protocol for total chemical protein synthesis by AFPS. Comparative analyses with proteins produced by biological expression validated that the synthetic proteins obtained in this study exhibited similar structure and function.

Overall, in addition to its impressive speed, this new technique leads to a three-fold increase in peptide size limit achievable by AFPS techniques. This advance should facilitate future protein synthesis endeavors and influence multiple fields of research. Coupling AFPS technology with existing NCL methods could deliver >300-residue proteins after a single ligation reaction. Moreover, it is easy to imagine

Fast-flow automated protein synthesis

Hartrampf *et al.* optimized concentration, flow rate, reagents, additives, and temperature to minimize unwanted by-products in peptide synthesis. The peptide lengths are in the range of those of single-domain proteins.



cycle to less than 1 min, and convenient inline ultraviolet-visible monitoring has allowed quantitative analysis of the fluorenylmethoxycarbonyl (Fmoc) deprotection reactions. This feature offers the marked advantage of identifying problematic coupling steps rapidly and consistently throughout the synthesis, versus the time-consuming manual analyses typically used.

In the present study, Hartrampf *et al.* have methodically optimized the parameters of their AFPS setup, including concentration, flow rate, reagents, additives, and temperature, to surmount the difficult task of reducing by-products sufficiently to enable protein synthesis (5). They established particular sets of optimal conditions for each amino acid incorporation by first using the 30-residue glucagon-like peptide-1 as a test. They then applied their optimized procedure to the synthesis of nine different proteins, thus establishing a generalized

similar optimization efforts for a wide variety of unnatural amino acid residues, expanding their utility far beyond their current scope and potentially yielding all-unnatural protein-mimetic material that is inaccessible by biological methods. In that respect, it will not only catch up to but surpass ribosome capabilities. ■

REFERENCES AND NOTES

1. S. B. H. Kent, *Chem. Soc. Rev.* **38**, 338 (2009).
2. V. Agouridas *et al.*, *Chem. Rev.* **119**, 7328 (2019).
3. R. B. Merrifield, *J. Am. Chem. Soc.* **85**, 2149 (1963).
4. P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **266**, 776 (1994).
5. N. Hartrampf *et al.*, *Science* **368**, 980 (2020).
6. R. Behrendt, P. White, J. Offer, *J. Pept. Sci.* **22**, 4 (2016).
7. L. Wang, P. G. Schultz, *Angew. Chem. Int. Ed.* **44**, 34 (2004).
8. A. J. Mijalis *et al.*, *Nat. Chem. Biol.* **13**, 464 (2017).

ACKNOWLEDGMENTS

The author gratefully acknowledges North Carolina State University for startup support.

10.1126/science.abb9711

CORONAVIRUS

The search for a COVID-19 animal model

A comparison of SARS-CoV-2 replication, transmission, and disease in mice to monkeys

By **Seema S. Lakdawala¹** and
Vineet D. Menachery²

As the pandemic caused by severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2) continues to cause worldwide upheaval, scientists are racing to find appropriate animal models to study the coronavirus disease 2019 (COVID-19) attributed to the virus. The optimal animal model will depend on the scientific question. On page 1016 of this issue, Shi *et al.* (1) describe severe viral burden and airborne transmission of SARS-CoV-2 between cats and ferrets, highlighting an important animal model for SARS-CoV-2 transmission. Additionally, on page 1012 of this issue, Rockx *et al.* (2) found that young and aged cynomolgus macaques infected with SARS-CoV-2 shed virus in the upper and lower respiratory tract, but failed to develop severe clinical symptoms. These animal models offer distinct platforms to ask specific questions about SARS-CoV-2 infection, induction of disease, and transmission.

Humans with COVID-19 display a wide range of disease symptoms, from asymptomatic to severe pneumonia (3). Translating data from a single animal model to the varied disease outcomes in humans is not only challenging, but potentially misleading. Studies examining the efficacy of vaccines and antiviral drugs traditionally use models of severe disease, which may not mimic the common pathology in the majority of COVID-19 patients and could limit understanding of other important questions, including infection dynamics and transmission. Previous work on other emerging coronaviruses, such as Middle East respiratory syndrome-coronavirus (MERS-CoV) and SARS-CoV-1, have included mice, hamsters, ferrets, monkeys, and camels as animal models (4). Although mice are the preferred research animal model because of their cost, reproduction rate, and wealth of reagents available for

studying this species, early reports indicate that they are unsuitable for SARS-CoV-2 infection, likely due to receptor incompatibility (5). Therefore, transgenic mice that express the human angiotensin-converting enzyme 2 (hACE2), which is the host cell receptor for SARS-CoV-2 entry, will be useful to examine COVID-19 (5–7). Preliminary work with hACE2 mice demonstrates susceptibility but limited disease severity (8) (see the table).

Shi *et al.* administered a large dose of SARS-CoV-2 intranasally to a wide range of animals—ferrets, cats, dogs, pigs, chickens, and ducks—to test replication, pathogenesis, and transmission. The virus replicated

a combined intranasal and intratracheal administration shed virus in the upper and lower respiratory tract, but clinical symptoms were mild. Similarly, Gao *et al.* (9) tested the efficacy of an inactivated vaccine in rhesus macaques. In this study, vaccination with inactivated SARS-CoV-2 produced antibodies against the viral spike protein and nucleoprotein in mice, rats, and macaques. Macaques vaccinated with 3 or 6 µg of inactivated virus in alum were challenged 22 days later and demonstrated reduced viral RNA in nasal and anal swabs.

Together, these studies, along with preliminary studies in hamsters (10), suggest that there are only a handful of susceptible

Searching for the best animal model to study COVID-19

Comparison of currently available animal models for SARS-CoV-2 infection and COVID-19.

ANIMAL MODEL	UPPER RESPIRATORY TRACT*	LOWER RESPIRATORY TRACT†	FECES/FECAL SWAB	CONTACT TRANSMISSION	AIRBORNE TRANSMISSION	WEIGHT LOSS	SOURCE
Cat (6 to 9 months)	Y‡	N	Y	NR	Y (33%)	NR	(1)
Chicken	N	N	NR	N	NR	NR	(1)
Dog	N	N	Y	N	NR	NR	(1)
Duck	N	N	NR	N	NR	NR	(1)
Ferret	Y‡	Y§	Y	Y (100%)	Y (30%)	NR	(1, 14)
hACE2 mouse	NR	Y‡	Y‡	NR	NR	Y	(8)
Hamster	Y‡	Y‡	Y‡	Y (100%)	NR	Y	(10)
Kitten	Y‡	Y‡	NR	NR	Y (33%)	NR	(1)
Macaque	Y	Y	NR	NR	NR	N	(2, 9)
Pig	N	N	NR	N	NR	NR	(1)

* Includes nasal washes, nasal swabs, nasal turbinate, soft palate, trachea. † Includes lung sections. ‡ Denotes detection of infectious virus.

§ Infectious titer in the ferret lung was only reported in (14). COVID-19, coronavirus disease 2019; hACE2, human angiotensin-converting enzyme 2; NR, not reported; SARS-CoV-2, severe acute respiratory syndrome-coronavirus 2.

efficiently in the upper respiratory tract of cats and ferrets, suggesting that these animals are permissive to SARS-CoV-2. However, no severe clinical symptoms such as weight loss or respiratory distress were noted in ferrets or cats. In the other models, virus was not detected in nasal or rectal swabs, with the exception of two dogs in which SARS-CoV-2 RNA was detected in rectal swabs and the animals produced antibodies against the virus. These data suggest that pigs, chickens, and ducks are not permissive to SARS-CoV-2 infection.

Rockx *et al.* found that cynomolgus macaques infected with SARS-CoV-2 using

SARS-CoV-2 animal models (hamsters, ferrets, cats, and nonhuman primates). Susceptibility is likely driven by either binding affinity to the host ACE2 receptor or by differences in host protease activity on the spike protein, both key factors in coronavirus cell entry (11). In addition to release of infectious virus in the respiratory tract, viral RNA was found in rectal swabs of cats, ferrets, and macaques (1, 2), which is consistent with initial clinical observations of patients with COVID-19 (12) and suggests a potential role for the fecal-oral transmission route.

Epidemiological data have clearly de-

¹Department of Microbiology and Molecular Genetics, Center for Vaccine Research, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. ²Department of Microbiology and Immunology, Institute for Human Infection and Immunity, World Reference Center for Emerging Viruses and Arboviruses, University of Texas Medical Branch at Galveston, Galveston, TX, USA. Email: lakdawala@pitt.edu; vimenach@utmb.edu

fined that COVID-19 has a higher case fatality rate in individuals over 60 years of age (3). Yet, most in vivo research models use young, healthy animals. Models that recapitulate this aging phenotype would be incredibly useful to examine the underlying biology behind this phenomenon. To address this, Shi *et al.* examined COVID-19 symptoms in kittens versus cats aged 6 to 9 months, but found the opposite phenotype. COVID-19 was more severe in kittens; one kitten died on day 3 after infection and had a larger viral distribution compared to older cats. This variation in viral dissemination in kittens may be reminiscent of the vast tissue tropism and variation in disease severity exhibited by SARS-CoV-2 in humans (12). By contrast, Rockx *et al.* tested SARS-CoV-2 in young and aged macaques and did not observe any age-dependent differences. Of note, the virus is not consistently lethal in any of the animals tested thus far, nor does SARS-CoV-2 infection in these animals recapitulate the severe clinical symptoms observed in humans. As these animal models continue to be developed, attention should be paid to the role of age and other health conditions; these factors may be critical parameters that are necessary to fully evaluate human disease.

Transmission of viruses between people, either through contact (direct or indirect) or virus-containing aerosols, is a key determinant of viral disease burden globally. An important aspect of the SARS-CoV-2 pandemic, and previous influenza pandemics, is efficient airborne transmission of the virus. Airborne transmission can include a wide range of aerosol sizes. At close contact ranges, the exposure to large and small aerosols containing viruses is high. Shi *et al.* examined airborne transmission of SARS-CoV-2 between kittens and aged cats and observed that in both scenarios, the virus could transmit through the air to 33% of naïve recipient cats or kittens. In these studies, animals were separated by perforated barriers that limit physical contact but allow for air to be shared between the experimentally infected donor and the susceptible recipient. Studies of influenza virus transmission have indicated that viral replication in the upper respiratory tract, and specifically the soft palate, play an important role in airborne transmission (13). Shi *et al.* found that SARS-CoV-2 replicated in the soft palate of cats, kittens, and ferrets. Although ferret transmission was not examined in this study, a report suggested a similar airborne transmission rate of 30% for SARS-CoV-2 in ferrets (14). No contact transmission between dogs and other animals (pigs, ducks, and chickens) was observed (1).

The transmission of SARS-CoV-2 between cats highlights the susceptibility of this animal model to infection. Consistent with this observation, transmission of SARS-CoV-2 from humans to tigers was recently documented, as was virus spread among big cat units in the Bronx Zoo (15). On the basis of data from Shi *et al.*, infected cats appear asymptomatic, so infections in cats may go undetected. Additional studies are needed into the seroprevalence of SARS-CoV-2-specific antibodies in cats and identification of coronaviruses from this animal source to ascertain the potential for cats to be an intermediate host for SARS-CoV-2.

As the pursuit of SARS-CoV-2 vaccines and antivirals surges on, animal models play the most important role to determine the effectiveness of potential therapeutic strategies. The available studies suggest that hamsters, ferrets, and cats may serve as attractive alternatives to nonhuman primate and transgenic mouse studies. Because hamsters and transgenic mice display the most severe clinical symptoms, such as weight loss, they may provide robust small-animal models for studying efficacy of various vaccine platforms. By contrast, cats and ferrets may provide a useful model system for studying transmissibility of the virus and the effectiveness of antivirals to limit spread. With robust reduction in viral load as presented by Gao *et al.* (9), nonhuman primates may offer the most relevant model to assess vaccine and antiviral effectiveness before rapid deployment to humans. Therefore, continued evaluation of mice to nonhuman primate models will provide critical data on the animals best suited to study the many open questions about COVID-19. ■

REFERENCES AND NOTES

1. J. Shi *et al.*, *Science* **368**, 1016 (2020).
2. B. Rockx *et al.*, *Science* **368**, 1012 (2020).
3. Z. Wu, J. M. McGoogan, *JAMA* **323**, 1239 (2020).
4. L. M. Gretebeck *et al.*, *Curr. Opin. Virol.* **13**, 123 (2015).
5. P. Zhou *et al.*, *Nature* **579**, 270 (2020).
6. P. B. McCray Jr. *et al.*, *J. Virol.* **81**, 813 (2007).
7. V. D. Menachery *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3048 (2016).
8. L. Bao *et al.*, *bioRxiv* 10.1101/2020.02.07.939389 (2020).
9. Q. Gao *et al.*, *Science* 10.1126/science.abc1932 (2020).
10. S. F. Sia *et al.*, *Research Square* 10.21203/rs.3.rs-20774/v1 (2020).
11. V. D. Menachery *et al.*, *J. Virol.* **94**, (2020).
12. Y. Chen *et al.*, *J. Med. Virol.* 10.1002/jmv.25825 (2020).
13. S. S. Lakdawala *et al.*, *Nature* **526**, 122 (2015).
14. Y. I. Kim *et al.*, *Cell Host Microbe* **20**, 30187 (2020).
15. N. Daly, "Seven more big cats test positive for coronavirus at Bronx Zoo," *National Geographic*, 22 April 2020.

ACKNOWLEDGMENTS

The authors are funded by NIH CEIRS HHSN272201400007C (S.S.L.) and R00 AG049092 (V.D.M.).

10.1126/science.abc6141

IMMUNOLOGY

Killer cells add fire to fuel immunotherapy

Cytotoxic lymphocytes induce inflammatory target cell death, amplifying antitumor immunity

By Christopher J. Nicolai and David H. Raulet

Cytotoxic lymphocytes, including natural killer (NK) cells and cytotoxic T lymphocytes (CTLs or CD8⁺ T cells), mediate antiviral and antitumoral immunity. Target cell killing by CTLs and NK cells is primarily mediated through the release of cytotoxic granules that contain serine proteases called granzymes and a pore-forming protein, perforin. Perforin delivers granzyme B (GZMB) into target cells, where it initiates apoptosis, a noninflammatory form of programmed cell death (1, 2). On page 965 of this issue, Zhou *et al.* (3) report their discovery of a new mechanism of cytotoxicity in which granzyme A (GZMA), delivered to certain target cells by NK cells and CTLs, activates gasdermin B (GSDMB), a pore-forming protein, which causes a proinflammatory form of cell death called pyroptosis (4). Expression of GSDMB by mouse tumor cells conferred better tumor control in response to immune checkpoint therapy.

Programmed cell death can occur through a variety of pathways that result in distinct biological outcomes. The best studied programmed cell death pathway is apoptosis, in which cytoplasmic and nuclear material condense into membrane-bound fragments, called apoptotic bodies, which are phagocytosed and destroyed by macrophages (2). Cell death through apoptosis is considered noninflammatory and is necessary for organismal development and homeostasis. By contrast, pyroptosis ("fiery death") is a proinflammatory form of cell death, usually activated by multiprotein oligomers called inflammasomes, which are triggered by a variety of stimuli. Inflammasomes activate inflammatory caspase-1, -4, -5, or -11, leading to lytic cell death and the release of proinflammatory

Department of Molecular and Cell Biology,
University of California, Berkeley, Berkeley, CA, USA.
Email: raulet@berkeley.edu

molecules, such as the cytokines interleukin-1 β (IL-1 β) and IL-18, and eicosanoids (4, 5). Two pivotal studies identified gasdermin D (GSDMD) as a target of inflammatory caspases and the key executor of pyroptosis (6, 7). Upon cleavage, the amino-terminal domain of GSDMD oligomerizes to form pores in the cell membrane, leading to cell permeabilization and cytokine release. GSDMD is one of a family of gasdermin proteins (8) that have pore-forming domains but differ in the proteases that activate them. The roles of some members of the gasdermin family are unknown.

Zhou *et al.* provide compelling evidence that cytolysis by CTLs and NK cells is sometimes pyroptotic instead of apoptotic. In addition to GZMB, granules released by cytotoxic lymphocytes contain GZMA, which directly cleaves GSDMB in the cytosol, un-

to amplify. The results support the view that cytotoxic lymphocytes both kill target cells and in some cases activate additional inflammatory signals through pyroptosis that amplify the immune response.

GSDMB was more frequently expressed in human tumors that arose in mucosal and gastric tissues than in other types of tumors, although less often than in corresponding normal tissues, possibly reflecting selective loss of GSDMB in advanced cancers. In some types of human cancer, notably bladder carcinoma and skin cutaneous melanoma, there was a correlation between GSDMB expression and higher patient survival, although GSDMB expression in renal clear cell carcinomas was associated with lower patient survival.

Pyroptotic killing induced by cytotoxic lymphocytes has been observed previously,

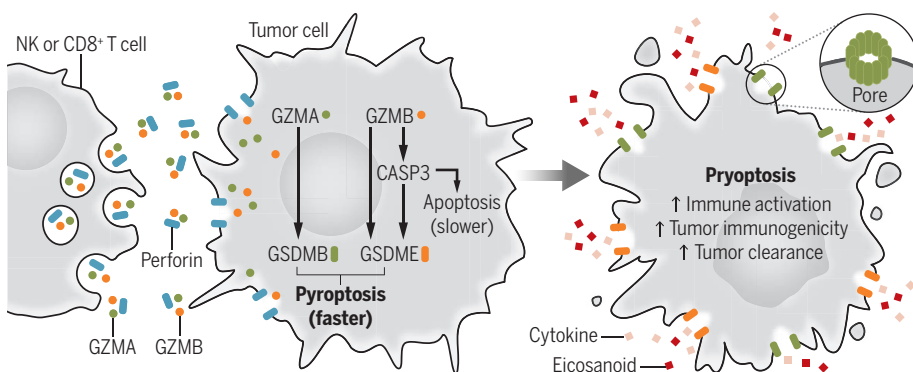
after CAR-T cell treatments in mouse models. The authors reported that GSDME-dependent pyroptosis was induced only when the T cells encountered the strong activating signals mediated by CAR engagement with CD19 and not through typical T cell antigen receptor-mediated signaling. Furthermore, another study showed that introducing nanoparticles containing active GSDMA into tumor cells mobilizes powerful antitumor T cell responses (11). It was shown that IL-1 β , and to a lesser extent IL-18, were necessary for this amplified antitumor response.

Together with the study of Zhou *et al.*, these studies suggest that pyroptosis is a common outcome of encounters between cytotoxic lymphocytes and susceptible target cells and that pyroptosis provides a feed-forward mechanism to amplify cellular immunity and tumor rejection (see the figure). These studies do not yet provide a detailed understanding of how pyroptosis amplifies the antitumor response. Inflammatory signaling coupled with the release of target cell antigens are likely important, but the relevant inflammatory signals remain poorly defined. In one case, IL-1 β and possibly IL-18 were shown to play a role in pyroptosis-induced immune enhancement (11), whereas another report argued that IL-1 β plays no role (9). If IL-1 β and/or IL-18 are important, it remains unclear exactly how they mediate their amplifying effects.

Moreover, it is not clear how the cytokines are elaborated in these circumstances. In inflammasome-induced pyroptosis, “priming” of cells by pattern-recognition receptors (which recognize pathogens) is often required for synthesis of pro-IL-1 β and pro-IL-18, and the cytokines are processed by caspases to generate the active forms (5). How priming might occur in tumor cells is unclear, but activation of stimulator of interferon genes (STING; a sensor of cytoplasmic DNA) is an intriguing possibility (12). Although much remains to be learned, these findings require a revision of how cell-mediated cytotoxicity occurs and its consequences. ■

Pyroptosis amplifies cellular immunity

Granzyme B (GZMB) and GZMA released by cytotoxic lymphocytes [natural killer (NK) and CD8 $^{+}$ T cells] penetrate target cells with the help of perforin. Recent evidence reveals that GZMA directly cleaves gasdermin B (GSDMB), whereas GSDME can be activated directly by GZMB but also indirectly by caspase 3 (CASP3). Activation of GSDMB or GSDME initiates pyroptosis, which promotes inflammation that amplifies cellular immunity.



leashing its amino-terminal pore-forming domain and initiating pyroptosis. Because this can be faster than apoptosis, cells that expressed GSDMB underwent pyroptosis rather than apoptosis. GSDMB expression was rare in a large panel of human tumor cell lines, but its expression could be induced in ~30% of the lines upon stimulation with cytokines, especially interferon- γ (IFN- γ), a pleiotropic cytokine secreted by activated lymphocytes, including CTLs and NK cells. Mice do not express a GSDMB homolog yet still express GZMA. Mice bearing tumors engineered to express GSDMB showed greater tumor rejection than did mice bearing GSDMB-deficient tumors when combined with anti-programmed cell death protein 1 (PD-1) immunotherapy. These findings suggest that PD-1 blockade was necessary to unleash antitumor immune responses that GZMA-dependent GSDMB activation helped

although mediated by other gasdermin family members. For example, GZMB from NK cells directly cleaved and activated GSDME in target cells and concomitantly activated caspase 3, which also activates GSDME (9). Cleaved GSDME initiated pyroptosis and amplified immune infiltration of mouse mammary tumors grafted in mice, macrophage phagocytosis of tumor cells, and tumor rejection. Disruption of the *Gsdme* gene in mouse tumor cell lines resulted in faster-growing tumors and reduced survival in vivo, building on previous evidence that GSDME is a tumor suppressor that is frequently mutated or repressed in tumors.

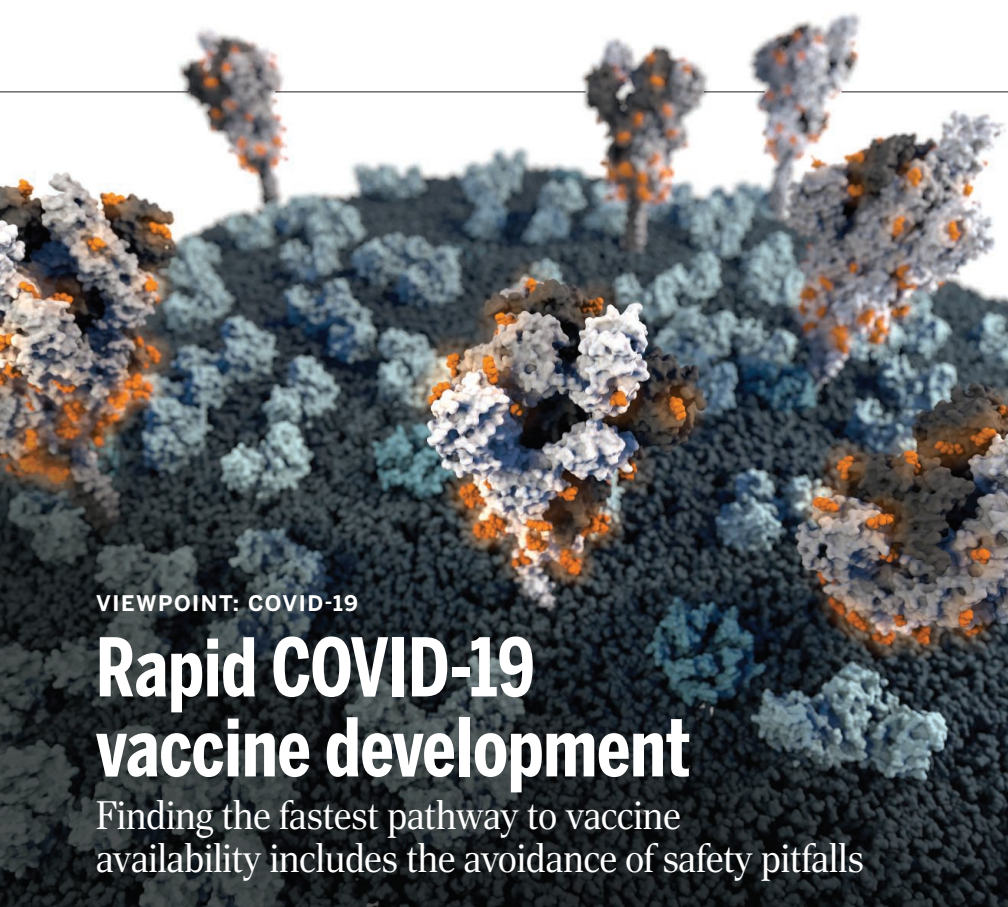
Killing of human CD19 $^{+}$ leukemia cells in vitro by CD19-targeting chimeric-antigen receptor (CAR) T cells initiates tumor cell pyroptosis (10). In this case, pyroptosis was associated with systemic inflammation, known as cytokine release syndrome,

REFERENCES AND NOTES

1. I. Voskoboinik *et al.*, *Nat. Rev. Immunol.* **15**, 388 (2015).
2. J. F. Kerr *et al.*, *Br. J. Cancer* **26**, 239 (1972).
3. Z. Zhou *et al.*, *Science* **368**, eaaz7548 (2020).
4. T. Bergsbaken *et al.*, *Nat. Rev. Microbiol.* **7**, 99 (2009).
5. K. Schroder, J. Tschopp, *Cell* **140**, 821 (2010).
6. J. Shi *et al.*, *Nature* **526**, 660 (2015).
7. N. Kayagaki *et al.*, *Nature* **526**, 666 (2015).
8. P. Broz *et al.*, *Nat. Rev. Immunol.* **20**, 143 (2020).
9. Z. Zhang *et al.*, *Nature* **579**, 415 (2020).
10. Y. Liu *et al.*, *Sci. Immunol.* **5**, eaax7969 (2020).
11. Q. Wang *et al.*, *Nature* **579**, 421 (2020).
12. S. S. Ho *et al.*, *Immunity* **44**, 1177 (2016).

ACKNOWLEDGMENTS

We thank R. Vance, K. Deets, and A. Tubbs for helpful comments.



VIEWPOINT: COVID-19

Rapid COVID-19 vaccine development

Finding the fastest pathway to vaccine availability includes the avoidance of safety pitfalls

By **Barney S. Graham**

Rapid development of a vaccine to prevent coronavirus disease 2019 (COVID-19) is a global imperative, and defining the stakes and potential hurdles is critical because regulatory and medical decisions are based on benefit:risk calculations. The ability of viruses to achieve pandemic spread is diminished by establishing higher levels of community (herd) immunity, and a key question is whether protection against severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2) will happen by widespread deployment of an effective vaccine or by repeated waves of infection over the next few years until ~60 to 70% of people develop immunity. Because the human population is naïve to SARS-CoV-2, the consequences of repeated epidemics will be unacceptably high mortality, severe economic disruption, and major adjustments to our way of life. Therefore, the benefit of developing an effective vaccine is very high, and even greater if it can be deployed in time to prevent repeated or continuous epidemics.

Vaccine development is usually measured in decades, so having access to approved vaccines available for large-scale distribution before the end of 2020 or even 2021

would be unprecedented. However, new manufacturing platforms, structure-based antigen design, computational biology, protein engineering, and gene synthesis have provided the tools to now make vaccines with speed and precision. Antiviral vaccines can be classified into two broad categories. Gene-based vaccines deliver gene sequences that encode protein antigens that are produced by host cells. These include live-virus vaccines, recombinant vaccine vectors, or nucleic acid vaccines. Protein-based vaccines include whole-inactivated virus, individual viral proteins or subdomains, or viral proteins assembled as particles, all of which are manufactured *in vitro*. Recombinant vaccine vectors and nucleic acid vaccines are best suited for speed because they can be more easily adapted to platform manufacturing technologies in which upstream supply chains and downstream processes are the same for each product. Precision is achieved by knowing the atomic structure of the vaccine antigen and preserving the epitopes targeted by the vaccine.

For any vaccine intended to generate antibody-mediated immunity, delivering a conformationally correct protein is critical. The CoV spike protein is displayed on the virus surface and carries out viral entry. It accomplishes this by undergoing a massive rearrangement that pulls the virus and cell membranes together and fuses them. Therefore, spike is a dynamic and metastable protein that has two major conformational

The trimeric spike protein is the primary target for vaccine-induced antibodies intended to block virus attachment to the human angiotensin-converting enzyme 2 (ACE2) receptor.

states, prefusion and postfusion. Displaying this antigen so that it maintains the surface contours and chemistry of the original native prefusion spike protein will preserve the epitopes required for eliciting high-quality neutralizing antibody responses. The vaccine formulation and delivery can also be crafted to influence T cell functions and response patterns. Gene-based delivery can induce CD8⁺ T cells and generally drive a CD4⁺ T helper 1 cell (T_H1)–type immune response, which has favorable antiviral properties. Adjuvants not only can be used to improve the magnitude and durability of antibody responses induced by protein-based vaccines but can also influence T cell–derived cytokine patterns and thus modulate immune responses.

Safety is a primary goal for vaccines that are given to otherwise healthy people, and there is a risk that vaccination could make subsequent SARS-CoV-2 infection more severe. This has happened before with vaccines based on whole-inactivated virus formulated in alum for a coronavirus of cats and for another unrelated respiratory virus in children. There are two different syndromes previously associated with vaccine-enhanced disease (see the table). One is antibody-dependent enhancement (ADE) (1) and the other is vaccine-associated enhanced respiratory disease (VAERD) (2). ADE is an Fc (the tail end of an antibody)–mediated enhancement of infection typically associated with flaviviruses, such as dengue virus (3). ADE is measured *in vitro* on cells that express Fc receptors (FcRs) either naturally or by transfection. The ADE mechanism involves increased binding efficiency of virus-antibody complexes to FcR-bearing cells, which triggers viral entry. This is more likely to occur when vaccine-induced antibody fails to effectively neutralize the virus because of insufficient concentration or affinity or the wrong specificity.

ADE has been described for feline infectious peritonitis virus (FIPV), a coronavirus that targets macrophages for infection (tropism) and causes a systemic vasculitis-like disease (4). Antibody-mediated disease enhancement was demonstrated after infection in cats that were previously vaccinated with alum-adsorbed inactivated virus. Although SARS-CoV-2 cellular tropism has not been completely defined, it is a respiratory virus, and consistent with coronaviruses that cause Middle East respiratory syndrome (MERS-CoV) and SARS (SARS-CoV-1), infection of respiratory epithelium results in a very different pathogenesis than the macrophage-tropic FIPV. ADE was shown to occur for SARS-

Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. Email: bgraham@mail.nih.gov

ILLUSTRATION: V. ALTOUNIAN/SCIENCE

CoV-1 in vitro in an FcR-bearing human B cell lymphoma cell line and based on detection of viral gene fragments by polymerase chain reaction (PCR). However, there is no experimental data in vivo showing that this type of antibody-mediated entry is relevant to the pathophysiology of respiratory coronaviruses like SARS-CoV-1, and even in vitro, no infectious virus was produced, suggesting an abortive replication cycle in myeloid cells (5).

VAERD is a distinct clinical syndrome that occurred in young children in the 1960s when whole-inactivated virus vaccines for measles and respiratory syncytial virus (RSV) were tested (6, 7). Immunizing with limiting doses of RSV antigen, especially with conformationally incorrect antigens, can result in two major types of immunological phenomenon that correlate with enhanced respiratory disease (ERD). One is a relatively high ratio of binding antibody to neutralizing antibody. Having a large amount of antibody that binds, but does not neutralize, virus in the presence of a high viral load could potentially result in immune complex deposition

clearance. Therefore, avoiding T_H2-biased immune responses may be important, especially in young infants with small airways that can be easily obstructed. In the youngest cohort of children who received the FI-RSV vaccine, 80% of those infected required hospitalization compared with 5% of placebo recipients (7). In the lung histopathology of the two children who died, there was an abundant polymorphonuclear leukocyte response in lungs that included eosinophils. This is consistent with findings in animal models of T_H2-biased CD4⁺ T cell responses associated with FI-RSV and VAERD (11).

Similar T cell and cytokine response patterns have been shown in mice, cotton rats, cattle, and nonhuman primate models of RSV immunized with whole-inactivated virus formulated with alum (12). In almost all the examples of VAERD demonstrated in humans for RSV and measles, and in animals for SARS (13), the vaccine antigen was whole-inactivated virus. There are caveats to the animal experiments because the allergic inflammation phenomenon can also be elic-

T_H2-biased immune responses. Using limiting dilutions of vaccines and examining lung pathology in animals with breakthrough infection after challenge should also help gauge the likelihood of aberrant pathology in vaccinated humans.

Defining the immunological parameters of VAERD in animal models of RSV delayed extensive industry involvement in vaccine development by ~30 years. Although the potential risk of vaccine-induced antibody or T cell responses leading to adverse responses to natural SARS-CoV-2 infection should be carefully evaluated, there is also a risk of delaying clinical trials in favor of prolonged evaluation of vaccines in animal models that do not fully recapitulate the pathogenesis of disease in humans. In the midst of a pandemic, it is reasonable to require certain qualities in candidate vaccines as described above and to start phase 1 clinical trials based on preliminary immunogenicity in animals and expanded trials based on human immunogenicity and evidence of protection in animal models. Justifying expansion to thousands of subjects in efficacy trials (that is, phase 2 and 3 trials) could include additional evidence of vaccine safety in animals immunized with limiting doses of vaccine and breakthrough infections after SARS-CoV-2 challenge. Judicious evaluation of candidate vaccines in healthy adults in parallel with vaccine studies in animal models and coincident process development to scale-up production capacity provides a path forward with minimal risk to human subjects and the potential for enormous benefit through accelerated COVID-19 vaccine availability. ■

Potential risks associated with vaccine development for COVID-19

Antibodies that bind virus without neutralizing infectivity can cause disease through increased viral replication or formation of immune complexes that deposit in tissue and activate complement pathways associated with inflammation. T helper 2 cell (T_H2)-biased responses have also been associated with ineffective vaccines that lead to enhanced disease after subsequent infection. Antibody-dependent enhancement (ADE) of viral replication has occurred in viruses with innate macrophage tropism. Virus-antibody immune complexes and T_H2-biased responses can both occur in vaccine-associated enhanced respiratory disease (VAERD).

Antibody-mediated			T cell-mediated
	ADE	VAERD	VAERD
Mechanism	Fc-mediated increase in viral entry	Immune complex formation and complement deposition	T _H 2-biased immune response
Effectors	Macrophage activation and inflammatory cytokines	Complement activation and inflammatory cytokines	Allergic inflammation and T _H 2 cytokines
Mitigation	Conformationally correct antigens and high-quality neutralizing antibody		T _H 1-biasing immunization and CD8 ⁺ T cells

and complement activation. This was demonstrated in the small airways of infants during the formalin-inactivated (FI) RSV vaccine trial in 1966 and contributed to inflammation and airway obstruction (8). A similar phenomenon occurred after measles infection of Rhesus macaques that were immunized with whole-inactivated measles virus vaccine (9).

The other observation is that immunization with whole-inactivated virus vaccines followed by RSV infection can result in allergic inflammation (10). Responses that accentuate production of the cytokines interleukin-4 (IL-4), IL-5, and IL-13 result in increased mucus production, eosinophil recruitment, airway hyperresponsiveness, and attenuated cytolytic T cell activity, collectively known as T_H2 immune responses. These events potentiate airway dysfunction and delay viral

ity by using the same cell line and media to grow the vaccine virus and challenge virus. Cellular components and media additives can cause sensitization to those proteins even without viral antigens present (14).

There are ways to mitigate the risks of vaccine-enhanced disease syndromes informed by prior work on RSV vaccines that should be considered for COVID-19 vaccine development (15). It will be important to demonstrate the potential for vaccine efficacy in early-phase clinical studies by measuring the induction of neutralizing antibodies and in animal models by demonstrating protection against virus replication and disease. Equally important will be using conformationally correct antigens to elicit high-quality, functionally relevant antibody and to avoid induction of non-neutralizing antibody and

REFERENCES AND NOTES

1. J. R. Mascola et al., *AIDS Res. Hum. Retroviruses* **9**, 1175 (1993).
2. F. P. Polack, *Pediatr. Res.* **62**, 111 (2007).
3. K. A. Dowd, T. C. Pierson, *Virology* **411**, 306 (2011).
4. C. W. Olsen, W. V. Corapi, C. K. Ngichabe, J. D. Baines, F. W. Scott, *J. Virol.* **66**, 956 (1992).
5. M. Jaume et al., *J. Virol.* **85**, 10582 (2011).
6. V. A. Fulginiti, J. J. Eller, A. W. Downie, C. H. Kempe, *JAMA* **202**, 1075 (1967).
7. H. W. Kim et al., *Am. J. Epidemiol.* **89**, 422 (1969).
8. F. P. Polack et al., *J. Exp. Med.* **196**, 859 (2002).
9. F. P. Polack, S. J. Hoffman, G. Crujeiras, D. E. Griffin, *Nat. Med.* **9**, 1209 (2003).
10. B. S. Graham et al., *J. Immunol.* **151**, 2032 (1993).
11. T. J. Ruckwardt, K. M. Morabito, B. S. Graham, *Immunity* **51**, 429 (2019).
12. A. Muralidharan, C. Li, L. Wang, X. Li, *Expert Rev. Vaccines* **16**, 351 (2017).
13. M. Bolles et al., *J. Virol.* **85**, 12201 (2011).
14. C. A. Shaw et al., *Vaccine* **31**, 306 (2013).
15. S. K. Browne, J. A. Beeler, J. N. Roberts, *Vaccine* **38**, 101 (2020).

ACKNOWLEDGMENTS

Thanks to J. Mascola, K. Morabito, K. Corbett, K. Bok, R. Seder, and M. Kanekiyo for helpful comments and discussions. This work was supported by intramural funding from the National Institute of Allergy and Infectious Diseases, NIH. The author is an inventor on patent applications that involve coronavirus vaccines and monoclonal antibodies.

Eleanor Margaret Burbidge (1919–2020)

Astronomer, astrophysicist, and champion of gender equality

By Alec Boksenberg

World-celebrated observational astronomer and astrophysicist Eleanor Margaret Burbidge (known as Margaret) died on 5 April at age 100. Margaret was a brilliant researcher, innovator, leader, and inspiration to others. She greatly advanced knowledge of the properties of stars and distant galaxies and, most famously, demonstrated with colleagues how almost all elements are fabricated from hydrogen within stars. By both practice and example, she was a force for gender equality in science and antidiscrimination in all forms. She served as the first female director of key scientific bodies and institutions in the United States and the United Kingdom.

Margaret was born in 1919 in Davenport, England. By the age of 12, she was consuming substantial books on astronomy. She graduated from University College London (UCL) in 1939 with a degree in astronomy with concentrations in mathematics and physics. Particularly attracted by the physical complexity of stars, Margaret enrolled as a graduate student at UCL while looking after the telescope instruments of the University of London Observatory at Mill Hill. During World War II, she worked on spectroscopic stellar research, taking advantage of the unusually dark skies created by blackouts and often sheltering from nearby German bomb blasts. She finished her Ph.D. in astronomy and astrophysics in 1943 and published her first scientific paper in 1946. In 1947, she met her husband to be, Geoffrey Burbidge, when he enrolled in the physics Ph.D. program at UCL. They married in 1948, soon after she inspired him to turn to astrophysics by inviting him to the observatory where she had become acting director.

They moved in 1950 to the United States in search of telescopes and sites superior to those in the United Kingdom. In response to her first job application, Margaret received a letter saying the position was not open to women. In 1951, she was welcomed by the University of Chicago's Yerkes Observatory in Wisconsin. Later, she applied to observe at

the Mount Wilson Observatory in California but was denied, so Geoffrey applied instead. Margaret succeeded in using the observatory through subterfuge—she posed as Geoffrey's assistant, whereas the reverse was true.

On an extended visit to the Institute of Astronomy in Cambridge, England, in 1955 to 1956, Margaret—who had already made many substantial observational discoveries—played a crucial role in the intensive work that produced one of the most important discoveries in the history of astrophysical research to this day. Along with Geoffrey, physicist William Fowler, and astrophysicist Fred Hoyle, Margaret demonstrated in extraordinary breadth and depth that the bulk of all elements in the Universe were synthesized within stars through progressive stages of nuclear fusion and final outburst.



Along with Fred Hoyle and Geoffrey, Margaret favored the steady-state theory of the Universe, which postulates that the Universe is constantly generating new matter to counter its own expansion, as opposed to the Big Bang theory, which states that there was a definite beginning. She contributed evidence for the steady-state theory through her observations, although her important work on the then-mysterious fast-moving quasars (now recognized to be distant galaxies powered by immense central black holes) suggested a different explanation. A chance discovery by others led to wide acceptance of the Big Bang theory, but Margaret's persistent radical work should be celebrated as revolutionary thinking that spurred efforts toward further observations and inspired new theories to progress to a conclusion.

Margaret was well known for her work opposing professional discrimination against

women. Perhaps remembering the cunning strategy required to conduct her own early observatory work, she often advised women to be relentless in finding alternative routes when something obstructed their path to a goal. In 1971, she declined the American Astronomical Society's influential Annie Jump Cannon Award, given exclusively to women, because she felt that it was itself discriminatory. She held many top leadership and administrative posts, including director of the Royal Greenwich Observatory in the United Kingdom from 1973 to 1975, president of the American Astronomical Society from 1976 to 1978, and president of the American Association for the Advancement of Science (the publisher of *Science*) in 1983. From 1979 to 1988, she was the first director of the Center for Astrophysics and Space

Sciences at the University of California, San Diego, where she worked from 1962 until her retirement. She was inducted into the Women's Museum of California Hall of Fame in 2003. Asteroid (5490) Burbidge was named after her.

I got to know Margaret through small scientific social gatherings in Cambridge and London, at which we exposed, explored, and debated ideas, sometimes vehemently but always in good humor. Margaret was calm, soft-spoken, and polite. Once, while discussing the Hubble Space Telescope Faint Object Spectrograph that she had helped develop, she showed me an old photographic record of a faint spectrum she had obtained from the ground. At first, I could see nothing on it except for a very faint, grainy haze. Then, with a broad smile, she tilted the photograph on its long side, visually condensing the image vertically and bringing the spectrum well into view. We both laughed at how creative thinking could provide clarity.

Throughout Margaret's life, she showed that in what was (and remains) a so-called man's world, women can and should be recognized for their brilliance and rise to the greatest heights of achievement and fame. Margaret did so with graciousness, humanity, and a drive to help others along the way. She made a difference not only to her field but to all scientists who can look to her as an example of persisting despite obstacles. ■

Institute of Astronomy, University of Cambridge,
Cambridge, UK. Email: boksy@ast.cam.ac.uk

10.1126/science.abc6555



VACCINATION: COVID-19

A strategic approach to COVID-19 vaccine R&D

A public-private partnership and platform for harmonized clinical trials aims to accelerate licensure and distribution

By **Lawrence Corey**^{1,2}, **John R. Mascola**³,
Anthony S. Fauci⁴, **Francis S. Collins**⁵

There is an unprecedented need to manufacture and distribute enough safe and effective vaccine to immunize an extraordinarily large number of individuals in order to protect the entire global community from the continued threat of morbidity and mortality from severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2). The global need for vaccine and the wide geographic diversity of the pandemic require more than one effective vaccine approach. Collaboration will be essential among biotechnology and pharma-

ceutical companies, many of which are bringing forward a variety of vaccine approaches (1). The full development pathway for an effective vaccine for SARS-CoV-2 will require that industry, government, and academia collaborate in unprecedented ways, each adding their individual strengths. We discuss one such collaborative program that has recently emerged: the ACTIV (Accelerating COVID-19 Therapeutic Interventions and Vaccines) public-private partnership. Spearheaded by the U.S. National Institutes of Health (NIH), this effort brings together the strengths of all sectors at this time of global urgency. We further discuss a collaborative platform for conducting harmonized, randomized controlled vaccine efficacy trials. This mechanism aims to generate essential safety and efficacy data for several candidate vaccines in parallel, so as to accelerate the licensure and distribution of multiple vaccine platforms and vaccines to protect against COVID-19 (coronavirus disease 2019).

We currently know little about what constitutes a protective immune response

against COVID-19. Data from SARS-CoV-1 patients as well as recently infected SARS-CoV-2 patients document relatively high levels of immune responses after infection, especially antibody responses to the surface (spike) protein that mediates entry into host cells. However, in vivo data on the type or level of immunity required to protect from subsequent re-infection, and the likely duration of that protection, are currently unknown. In animal models of SARS-CoV-1, immunization with recombinant subunit proteins and viral- and nucleic acid-vectored vaccines, as well as passive transfer of neutralizing antibodies to the spike protein, have been shown to be protective against experimental infection (2, 3). Endpoints vary from protection of infection to modification of viral replication and disease. These data bring optimism that a highly immunogenic vaccine will elicit the magnitude and quality of antibody responses required for protection. The role that T cell immunity plays in preventing acquisition or amelioration of early disease, either in animal challenge models or in human coronavirus disease, is unclear (4); this constitutes another reason why a diversity of vaccine approaches must be pursued.

A high degree of safety is a primary goal for any widely used vaccine, and there is theoretical risk that vaccination could make subsequent SARS-CoV-2 infection more severe. This has been reported for feline coronaviruses and has been observed in some vaccine-challenge animal models of SARS-CoV-1 (5). These preclinical data suggest that the syndrome of vaccine-associated enhanced respi-

¹Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA.

²Departments of Medicine and Lab Medicine, University of Washington, Seattle, WA 98195, USA. ³Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA. ⁴National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA. ⁵National Institutes of Health, Bethesda, MD 20892, USA. Email: afauci@niaid.nih.gov

ratory disease results from a combination of poorly protective antibodies that produce immune complex deposition together with a T helper cell 2 (T_H2)-biased immune response. The potential mechanism behind vaccine-induced immune enhancement and the means to minimize this risk have recently been reviewed (6). It will be important to construct conformationally correct antigens to elicit functionally effective antibodies—a lesson learned from vaccine-induced enhanced lower respiratory illness among infants receiving a formalin-inactivated respiratory syncytial virus (RSV) vaccine. Animal models of SARS-CoV-2 infection are currently being developed, and these models can be used to better understand the immune responses associated with protection (7).

CLINICAL AND IMMUNOLOGICAL ENDPOINTS

The primary endpoint for defining the effectiveness of a COVID vaccine also requires discussion. The two most commonly mentioned are (i) protection from infection as defined by seroconversion, and (ii) prevention of clinically symptomatic disease, especially amelioration of disease severity, including the frequency of disease requiring high-intensity medical care with some assessment of a decrease in hospitalization. This requires the close evaluation of the effect of vaccination on the severity of COVID-19 disease in a wide variety of epidemiological and medical settings among both younger and elderly populations as well as underserved minorities. All of these issues need to be evaluated in the context of these initial efficacy trials. Achieving these endpoints could also be associated with reduced transmissibility on a population basis.

Primary endpoints that involve reduction of disease require greater numbers of enrollees into trials, given that asymptomatic infection is estimated to be 20 to 40% of total cases of COVID-19 (8). Initial efficacy trials may then require a large initial enrollment, with ongoing monitoring of both serologic and clinical endpoints. A major challenge leading to a degree of complexity in developing clinical trial protocols for serological endpoints is the lack of precise knowledge of incidence rates (9). A critical requirement for such a multi-trial strategy is the establishment of independent laboratories with similar or identical validated serologic assays to provide a harmonizing bridge between multiple vaccine products and multiple vaccine efficacy trials. The use of these laboratories for each clinical trial, or the sharing of critical specimens from a trial, should be required. Parameters that would distinguish the immune response resulting from vaccination versus from infection are under intense in-

vestigation, and there is an immediate need to develop assays to address this issue.

Efficacy trials need to be evaluated for both benefit and harm. The likelihood of SARS-CoV-2 reexposure is much higher than that of SARS-CoV-1, which has disappeared from community circulation, and hence longer-term evaluation of potential enhancement with reexposure is needed. This requirement does not preclude licensure based on the endpoints outlined above; however, it does indicate that more prolonged follow-up of the initial vaccine cohorts should be undertaken. The durability of clinical and serologic endpoints will also need to be explored, as waning of immunity is common with human coronavirus infections (10). Coronaviruses have a single-stranded RNA genome with a relatively high mutation rate. Although there has been some genetic drift during the evolution of the SARS-CoV-2 epidemic, major alterations in the spike protein are not extensive to date, especially in the regions thought to be important for neutralization; this enables cautious optimism that vaccines designed now will be effective against circulating strains 6 to 12 months in the future (11).

The possibility of performing controlled human challenge trials, in which a small number of volunteers are vaccinated and subsequently challenged with SARS-CoV-2, has been suggested. Such experiments, if designed to define potential immune correlates or winnow out less effective vaccine approaches, may have utility. However, this approach has shortcomings with respect to pathophysiology and safety (12). Although the risk of severe disease or death in young healthy individuals from COVID-19 is quite low, it is not zero, and we do not yet have proven effective therapies for COVID-19 to rescue volunteers with complications from such a challenge. It is likely that a SARS-CoV-2 challenge strain will, by design, cause mild illness in most volunteers and thus may not recapitulate the pulmonary pathophysiology seen in some patients. Moreover, partial efficacy in young healthy adults does not predict similar effectiveness among older adults with major cofactors associated with COVID-19 disease, nor would it prove reduction of transmissibility to major susceptibility groups. Whether such experiments may be worthy of pursuit or would have a beneficial impact on timelines for vaccine development needs careful evaluation by an independent panel of ethicists, clinical trialists, and experts on vaccine development.

VACCINE PLATFORMS

It is encouraging that vaccine development efforts have moved swiftly, and several major vaccine platforms are moving toward clinical evaluation. These include traditional recom-

binant protein, replicating and nonreplicating viral vectors, and nucleic acid DNA and mRNA approaches. Each of these vaccine platforms has advantages and limitations. Important characteristics include speed and flexibility of manufacture, safety and reactogenicity, the profile of humoral and cellular immunogenicity, durability of immunity, scale and cost of manufacturing, vaccine stability, and cold chain requirements. No single vaccine or vaccine platform alone is likely to meet the global need, and so a strategic approach to the multi-pronged endeavor is absolutely critical.

Several companies are developing nucleic acid-based vaccines, including Moderna, BioNTech/Pfizer, CureVac (mRNA-based), and Inovio (DNA-based). DNA- and mRNA-based vaccines can be generated quickly on the basis of viral sequence, which allows a rapid pathway to the clinic (13, 14). Currently, optimal immunogenicity of DNA requires an electroporation or an injector delivery device to facilitate DNA entry into cells. mRNA vaccines use lipid nanoparticles to protect and deliver the mRNA and effectively adjuvant the immunogen. The scalability of these lipid nanoparticles and their temperature stability are issues that need to be addressed. Although there is a wide body of early-phase clinical experience with nucleic acid vaccines, none are licensed for widespread usage. As such, the path forward is filled with optimism, but some uncertainty remains, requiring rapid assessment of these products' immunogenicity and safety while addressing the lack of commercial experience with them.

Traditional recombinant protein technology can be used to express the spike protein (e.g., Sanofi, Novavax), and although the time to establish cell lines needed for manufacturing is longer than for nucleic acid vaccines, there is a robust commercial experience with protein and protein particle vaccines, including licensed vaccines for hepatitis B, human papillomavirus, varicella zoster, and influenza. Protein vaccines will require a potent adjuvant, which can be critical for inducing a predominantly T_H1 -type immune response; however, the availability of certain adjuvants may be limited. Viral vector vaccines encode the viral gene of interest into one of several well-characterized vectors, including adenovirus (Ad) and vesicular stomatitis virus (VSV). The replication-defective adenovirus 26 (rAd26), recently shown to be safe and immunogenic in preventing Ebola virus infection (15), is being developed by Janssen Pharmaceuticals for COVID-19. This platform has the potential to be manufactured at large scale. Preexisting immunity to the specific viral vector can attenuate immunogenicity, and this needs to be addressed in early-stage trials. A recombinant chimpanzee

Ad vector (ChAdOx1), developed by the University of Oxford and AstraZeneca, has also entered clinical trials. Similar versions of ChAd vaccine products have been tested in prior clinical trials and shown to be safe and immunogenic. The VSV vector vaccine platform is replication-competent and thus induces a robust, likely durable immune response with a single dose. A licensed VSV Ebola vaccine made by Merck is highly effective after a single dose, although its reactogenicity may be limiting in some populations. These diverse approaches provide the potential for scalable production required for widespread population use.

STRATEGIC COLLABORATIONS

Under the ACTIV public-private partnership, NIH has partnered with its sister agencies in the Department of Health and Human Services, including the Food and Drug Ad-

(DSMB) should be used so that the regulatory framework for the entire enterprise is coordinated and the regulatory agencies and the public can make objective assessment of the effect sizes between approaches. As vaccine candidates are poised to enter phase 1, the collective planning for phase 3 must be undertaken. Although much of this focus is on trials in the United States, the COVID-19 Prevention Networks established under the ACTIV program have a global focus, and coordination with the World Health Organization, Coalition for Epidemic Preparedness Innovations, and other global philanthropic partners must also occur.

Harmonized master protocols will be needed to enable transparent evaluation of the relative effectiveness of each vaccine approach. This harmonization can best be achieved through public-private partnerships such as ACTIV, in which government-

ment statisticians having access to cross-trial data in real time, and centralized immune monitoring laboratories. These innovations in the process of vaccine development are required to achieve the rapid development of the platform technologies entering clinical trials. Global effort, global cooperation, and transparency are needed to maximize the speed, veracity, and decision-making required to deliver scientific advances to the global population in a timely fashion. Models for all of these programs exist, and rapid implementation of these ideas is essential if we are to succeed in the timelines required to return us to pre-COVID-19 social interactions.

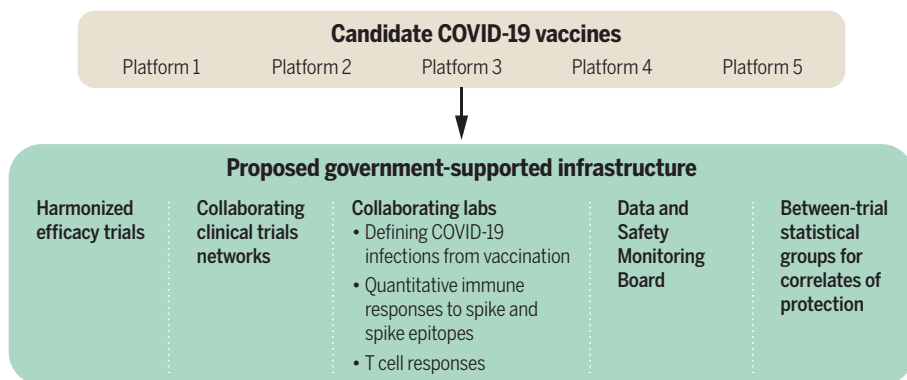
SCALE UP

The ability to manufacture hundreds of millions to billions of doses of vaccine requires the vaccine-manufacturing capacity of the entire world. Although new technologies and factories can be developed to sustain production, there is an immediate need to fund the necessary biomanufacturing infrastructure, including the fill/finish steps that provide vialled vaccine products for distribution. Cost, distribution system, cold chain requirements, and delivery of widespread coverage are all potential constriction points in the eventual delivery of vaccines to individuals and communities. All of these issues require global cooperation among organizations involved in health care delivery and economics.

To return to a semblance of previous normality, the development of SARS-CoV-2 vaccines is an absolute necessity. To achieve this goal, all the resources in the public, private, and philanthropic sectors need to participate in a strategic manner. The ACTIV public-private partnership and collaborative harmonized efficacy trials are enabling models to achieve our common goal. ■

The ACTIV model for SARS-CoV-2 vaccine development

The necessary partners in the public-private partnership are based on nonidentical but harmonized efficacy trials associated with collaborating clinical trials networks and laboratories, a common Data and Safety Monitoring Board, and an independent statistical group to determine correlates of protection.



ministration, Centers for Disease Control and Prevention, and Biomedical Advanced Research and Development Authority; other U.S. government departments including the Departments of Defense and Veterans Affairs; the European Medicines Agency; and representatives from academia, philanthropic organizations, more than 15 biopharmaceutical companies, and the Foundation for NIH. This forum allows for discussions and consensus on vaccine trial designs, rapid data sharing, and close collaborations between the public and private sectors to rapidly and efficiently conduct vaccine efficacy studies. There is an emerging consensus that vaccine trials need to either use common independent laboratories or contribute samples and data for the purpose of generating surrogate markers that ultimately speed licensure and an overall comparison of efficacy. A common Institutional Review Board as well as a common cross-trial Data and Safety Monitoring Board

supported central laboratories and independent biostatisticians serve as key resources for efficacy trials, thereby providing a standardized way to assess the relative immune responses of different types of vaccines (see the figure). Such laboratories enhance the ability to define correlates of protection, which would speed licensure for all vaccines as well as define populations that will achieve protective immunity. Data should be shared among companies and be provided to independent statistical evaluation, allowing the early evaluation of a potential surrogate marker of protection, which would markedly speed licensure and distribution. Such data can only be obtained from harmonization and collaboration early on, during the planning of efficacy trials and the implementation of the collaboration described in the figure: the use of collaborating clinical trial sites, the monitoring of these efficacy trials through a common DSMB, indepen-

REFERENCES AND NOTES

1. N. Lurie, M. Saville, R. Hatchett, J. Halton, *N. Engl. J. Med.* 10.1056/NEJMp2005630 (2020).
2. J. ter Meulen *et al.*, *Lancet* **363**, 2139 (2004).
3. H. Bisht *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 6641 (2004).
4. J. Huang *et al.*, *Vaccine* **25**, 6981 (2007).
5. M. Bolles *et al.*, *J. Virol.* **85**, 12201 (2011).
6. B. S. Graham, *Science* **368**, 945 (2020).
7. B. Rockx *et al.*, *Science* **368**, eabb7314 (2020).
8. T. M. McMichael *et al.*, *N. Engl. J. Med.* 10.1056/NEJMoa2005412 (2020).
9. S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, *Science* **368**, 860 (2020).
10. W. Liu *et al.*, *J. Infect. Dis.* **193**, 792 (2006).
11. T. Bedford *et al.*, medRxiv 2020.04.02.20051417 [preprint], 16 April 2020.
12. S. K. Shah *et al.*, *Science* **368**, 832 (2020).
13. K. A. Dowd *et al.*, *Science* **354**, 237 (2016).
14. N. Pardi *et al.*, *Nature* **543**, 248 (2017).
15. I. D. Milligan *et al.*, *JAMA* **315**, 1610 (2016).

ACKNOWLEDGMENTS

We thank C. Dieffenbach and E. Erbeling for insightful discussion and comments, and M. Miner for editorial assistance.

Published online 11 May 2020

10.1126/science.abc5312

sciencemag.org **SCIENCE**

Ethics and governance for digital disease surveillance

The question is not whether to use new data sources but how

By **Michelle M. Mello**^{1,2} and **C. Jason Wang**^{1,3}

Digital epidemiology—the use of data generated outside the public health system for disease surveillance—has been in use for more than a quarter century [see supplementary materials (SM)]. But several countries have taken digital epidemiology to the next level in responding to COVID-19. Focusing on core public health functions of case detection, contact tracing, and isolation and quarantine, we explore ethical concerns raised by digital technologies and new data sources in public health surveillance during epidemics. For example, some have voiced concern that trust and participation in such approaches may be unevenly distributed across society; others have raised privacy concerns. Yet counterbalancing such concerns is the argument that “sometimes it is unethical *not* to use available data” (1); some trade-offs may be not only ethically justifiable but ethically obligatory. The question is not whether to use new data sources—such as cellphones, wearables, video surveillance, social media, internet searches and news, and crowd-sourced symptom self-reports—but how.

INNOVATIONS AGAINST COVID-19

Some efforts involve escalations of existing techniques of digital epidemiology—for example, using cellphone signals and social media data to map the spread of the virus—whereas other, more innovative initiatives focus on implementing public health measures such as isolation and quarantine.

Disease modeling and forecasting using machine learning and artificial intelligence

There is growing potential to use machine learning and big data to forecast disease spread and prioritize people for testing or limitations on movement. One controversial application during the COVID-19 outbreak has been the Chinese government's require-

ment that citizens in more than 200 cities in stall an Alipay app on their smartphones that assigns a risk code to each person indicating the extent to which they are permitted to move around the community (see SM). The coding algorithm reportedly incorporates information on time spent at risky locations and frequency of contact with other people. Public dissatisfaction with the app arose from lack of transparency about the reasons

Whether to implement, how to implement

What is the alternative?

Whether to adopt a digital surveillance measure should be evaluated by referencing the counterfactual. What would be used instead of the technology, and is that more or less desirable?

Least burdensome alternative

For its use to be justified, a digital technology should be judged the least burdensome alternative that would accomplish the public health objective.

Public oversight

Oversight must include members of the public; focus on particular uses of the data, not just data transfers; and promote trustworthy, transparent, and convincing justifications for the decisions taken.

Correcting mistakes

Processes through which potential errors can be challenged in the courts may be modified, involving longer waits, less robust hearings, and reduced access to counsel. Proving mistakes can be difficult where a decision has been driven by an algorithm whose logic may not be transparent.

Inequities and bias

Inequities persist in people's access to the internet and cellphones. Even among those having access, disparities exist in who uses these technologies. These disparities risk creating bias in new data sets.

people were classified into particular groups and mismatch with individuals' own beliefs about their risk level. Yet algorithmic classification and prioritization of individuals or localities may offer an alternative to haphazard rollout of social distancing orders and COVID-19 testing.

Leveraging and linking large datasets for case identification

Governments have massive troves of citizens' personal data at their disposal that can be used to identify persons at increased risk of infection and prioritize them for investigation by health officials. The Taiwanese government linked immigration and customs data on travelers (in batch files, after deleting irrelevant travel history) to National Health Insurance data on hospital and clinic visits to identify individuals whose symptoms could be due to contracting the novel coronavirus during travel to an affected area (2). That information was shared with health care providers so that they could use it to make decisions during patient visits, such as asking for additional history of present illness and ordering a COVID-19 test.

Risk-based border security

Taiwan developed an interesting alternative to blanket travel restrictions: individualized risk assessment. Travelers scan a QR code using their smartphone, which leads to an online travel declaration form that asks for travel history and flight information, symptoms of fever or respiratory infection, and contact information in Taiwan. On the basis of their health and travel information, travelers are either sent a pass by text, asked to do home quarantine for 14 days, or instructed to self-isolate at home for 14 days (2).

Electronic monitoring of quarantined and isolated individuals

New Zealand, Thailand, and Taiwan use cellphone location data to monitor movement of persons subject to quarantine or isolation orders. In Taiwan, for example, violators can receive heavy fines or be ordered into facilities, but the government first messages individuals to instruct them to return home and asks local police to check on them (2). China, Poland, and Russia have gone further, using facial recognition software to monitor compliance with orders (see SM). Such measures, although intrusive, help reduce the need for labor-intensive, in-person monitoring. Deidentified location data from cellphones and social media apps can also be used to monitor population-level adherence to social distancing orders (see SM).

Digital technologies are also useful for supporting confined individuals. Remote monitoring through smartphones improves

¹Center for Health Policy/Primary Care and Outcomes Research, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA. ²Stanford Law School, Stanford, CA, USA. ³Department of Pediatrics and Center for Policy, Outcomes and Prevention, Stanford University School of Medicine, Stanford, CA, USA. Email: mmello@law.stanford.edu

the prospects for isolating and quarantining people at home rather than in facilities. People's temperatures can be transmitted by wearables, digital thermometers, or video, and health workers can regularly check on people's needs without exposing themselves to the risk of transmission. Communities, too, can mobilize to assist those confined at home, as is occurring in the United States through the neighborhood social networking app NextDoor.

Enhanced contact tracing

Serious doubts have been raised about whether traditional methods of contact tracing alone can arrest the COVID-19 epidemic. Adding algorithmic contact tracing through a cellphone app or operating system is proposed to reduce transmission of the virus through instantaneous notification of contacts (see SM) (3). A real-world experiment with such an approach is under way in Singapore, where in March 2020 the government requested that citizens install a government-developed smartphone app called TraceTogether. The app uses Bluetooth technology to exchange identifier numbers with the phones of other TraceTogether users within 6 feet of the user, sharing data with the government only if the user becomes subject to contact tracing because of a COVID-19 diagnosis (see SM). As of late April 2020, similar apps have been rolled out in nearly 30 countries, and a high-profile effort by Google and Apple to develop standards is under way (4).

The Israeli government has gone farther than Singapore, making use of infected persons' cellphone location data on an involuntary basis. Its approach sends texts to persons who come into contact with known COVID-19 cases to inform them that they must immediately quarantine themselves for 14 days (5). South Korea, too, has opted to use geolocation data without seeking consent. It publicly posts information on where infected persons traveled in the days before their diagnosis based on cellphone location data, credit card records, and surveillance video (5). No names are included, but individuals' age, nationality, and sex are. Taiwan used itineraries of passengers who disembarked the Diamond Princess cruise ship to send text alerts to people residing in areas the passengers visited, asking them to self-monitor and notify officials of any symptoms. The recipient list was compiled by using mobile phone base station positioning.

ETHICAL ISSUES RAISED

Ethical issues raised by digital epidemiology center on a core tension: these new uses of people's data can involve both personal and social harms, but so does failing to harness the enormous power of data to arrest epidemics.

Respecting privacy

Many epidemiologic uses of new data sources do not implicate informational privacy to a greater extent than current commercial and research practices—although some people may feel greater disquiet about governments using their data than they do companies or academics. Other uses involve larger potential intrusions on privacy. Using cellphone location and text data, in particular, goes beyond what citizens of democratic nations are accustomed to. Everyday uses of public and private data are typically not conducted with personal identifiers attached. Moreover, except for use by law enforcement, data are not ordinarily used for purposes of tracking down and imposing consequences on the subjects of the data. Contact tracing of the kind being carried out in Israel, by contrast, involves immediately imposing public health orders on those traced.

Respecting autonomy

Respect for individuals' autonomy generally requires asking them for permission to access their personal information and use it in particular ways. Informed consent is a bedrock principle of research ethics and medical care and is expressed—albeit weakly—in Terms and Conditions agreements for use of websites and apps, which ask users to agree to the company's planned uses of their data. The consent issue has particular salience for contact tracing through cellphone records because at least three alternative regimes—opt in, opt out, and mandatory—are possible, and different countries have made different choices.

Equity concerns

Use of new data sources can improve representation of some populations in epidemiologic analysis, including people who are underrepresented in data from laboratories and health care providers because they cannot or do not access care. Nevertheless, inequities persist across the globe in people's access to the internet and cellphones. Even in areas with access, disparities exist in who uses these technologies (see SM). These disparities risk creating bias in new data sets.

Minimizing the risk of error

The risk that governments will make errors in identifying areas and individuals at high risk of disease infection is heightened when using new data because of three factors: scope, speed, and sources. First, the use of large datasets means that a much greater number of people are under review than would ordinarily be the case; errors in even a small percentage of cases translate into large numbers of people affected. Second, the pressure to develop and roll out apps and al-

gorithms quickly during an emergency may mean compromises on testing and validation. But erroneously flagging individuals or areas can involve serious social and economic burdens, such as stay-at-home orders and business closures. Mistakes also undermine trust and waste limited public health resources (6). Third, the sources of information in some new datasets will be less reliable than traditional disease reporters. Particularly given the spread of misinformation about disease outbreaks through social media, the need to validate data derived from internet news, search data, and social media posts is acute. Self-reports of perceived symptoms, too, may be inaccurate or incomplete and are not easy to corroborate (see SM).

Correcting mistakes can pose special challenges during public health emergencies, underscoring the importance of taking steps up front to minimize the risk of error. Ordinary processes through which citizens and businesses can challenge public health orders in the courts may be modified, involving longer waits, less robust hearings, and reduced access to counsel. Proving mistakes can also be difficult when a decision has been driven by an algorithm because algorithmic logic is often not transparent.

Accountability

A key question is how to ensure that companies and governments conducting and using epidemiologic analyses of new data sources are accountable for what they do. Democratic processes ordinarily help ensure that policymaking is reasonably transparent, the public has opportunities for input, and irresponsible officials can be removed. But many initiatives during COVID-19 have been undertaken by countries without strong democratic traditions and free-speech protections. Even in the United States, technological solutions are being pursued by small groups of officials and tech company leaders working outside ordinary channels and public view. The need to make decisions quickly may justify such processes but increases concerns about responsible practices.

The potential for misappropriation of data collected and methods developed for disease surveillance looms large. After all, the same approaches that can be used for case identification and contact tracing can be used to identify and track a government's political opponents (5). Such fears undercut trust in what public health officials are trying to do, and without public trust and participation, many key strategies for fighting infectious disease cannot succeed.

POLICY RECOMMENDATIONS

Two principles should serve as lodestars when considering the ethics of digital surveil-

lance during pandemics. First, the wisdom of adopting a digital surveillance measure should be evaluated not in the abstract but by reference to the counterfactual. What would be used instead of the technology, and is that more or less desirable? The counterfactual for COVID-19 involves mass shelter-at-home and business closure orders, which impose serious liberty and economic deprivations and are, in most areas, completely nonconsensual. Digital surveillance offers the prospect of expediting the lifting of such orders and minimizing their use in future outbreaks (3). It may have particular value for vulnerable groups such as the elderly and persons with chronic illness who may otherwise remain confined after others are released.

The second principle is that for its use to be justified, a digital technology should be judged the least burdensome alternative that would accomplish the public health objective (7, 8). This principle has long driven thinking in public health ethics and law. For disease outbreaks, what constitutes the least restrictive alternative depends on the available public health resources, evidence concerning what behaviors people will engage in without coercive public health orders, features of the pathogen's transmissibility, and the stage of the epidemic. Even if such a weighing points toward the imposition of digital surveillance, the least restrictive alternative principle can help minimize privacy intrusions—for example, through data minimization (identifying the narrowest possible set of data elements, especially identifiable ones, and the minimum duration and scope of use required to achieve the objective) (8). We consider applications of these two principles to particular technologies with value in combating the novel coronavirus and similar pathogens.

Using algorithms in disease modeling and risk classifications

The use of artificial intelligence techniques for disease forecasting raises minimal ethical concerns if it uses individually deidentified data. But using personally identifiable information in algorithms that assign risk scores or categories to individuals, such as in the case of the Alipay app, must be considered more seriously because of the social consequences attached to these determinations. Here, the main concerns are the usual ones about algorithmic bias and error, and solutions offered for such problems in other contexts are applicable. These include making code and datasets publicly accessible and subject to peer review and continuing to refine the model as additional data become available. Additional safeguards should include creating mecha-

nisms for individuals to challenge algorithmic classifications, making classifications time limited, and using classifications to support recommendations rather than legal restrictions on individuals' movement.

Using electronic monitoring to support confined persons

Wider use of electronic monitoring to support individuals under confinement orders (isolation, quarantine, or shelter at home) should be pursued. It is aligned with both public health goals (because supported individuals are more likely to be able to stay home) and the principles of solidarity and reciprocity, which recognize societal obligations to support those called on to sacrifice liberty to prevent harm to others. Electronic monitoring is less intrusive than in-person visits by public health workers, safer for workers, and easier to scale, enabling officials



to reach more people. For diseases that are only transmissible when the infected person is symptomatic, virtually observed symptom checks may speed individuals' release from confinement by quickly ascertaining when they no longer present a danger to others. Public health officials should move quickly to expand use of virtual check-ins with persons confined at home, prioritizing those most likely to need assistance because of infection status, membership in a vulnerable group, and lack of social supports.

Using electronic monitoring to enforce restrictions on movement

The use of electronic monitoring to enforce confinement orders and travel restrictions is more problematic. In the United States, for example, the Supreme Court has held that a judicial warrant must be obtained, after showing probable cause to believe a person violated the law, to search private cellphone

records for law-enforcement purposes (see SM). It is unclear how courts would apply that precedent to enforcement of public health orders, but violation of some such orders is a crime.

Electronic monitoring by use of cellphone Bluetooth data, bracelets, or video cameras would likely be more effective in detecting public health order violations than current methods, which rely on police detection or police response to complaints. The question is whether more enforcement is better. The benefits of stringently enforcing mass shelter-at-home orders are not entirely clear, and the potential for strict enforcement—particularly through electronic eyes—to undermine trust in government and stoke resistance is troublesome.

The principles of the least restrictive alternative and proportionality (7) will be helpful in determining whether to use electronic

monitoring for public health orders in future disease outbreaks. There is a strong argument in favor of monitoring bracelets for persons who present a high risk of harm to others if officials have a reasonable suspicion that they will not comply with home isolation. It is less restrictive than the likely alternative for such individuals: confining them in secure facilities. Even in this context, electronic monitoring should not immediately trigger law-enforcement action; rather, the first intervention should be outreach by public health officials. They should seek to understand the reasons for a person's noncompliance, which may include misunderstanding of the order or inability to satisfy basic needs such as food and health care.

The justification for electronic enforcement attenuates as the gravity of potential harm and likelihood of noncompliance shifts. For these reasons, population-wide shelter-at-home orders present a weak case for using electronic surveillance for individual enforcement purposes. However, using electronic data to understand the extent of compliance at the population level—as is now being done by using deidentified cellphone data to measure travel distances from home—is well justified. Such information involves little or no privacy intrusion and has helped public health officers understand whether they need to issue clarifications of what is permitted and prohibited, tighten social distancing orders further, or provide additional economic and social supports to enable people to stay home.

Using cellphone data for contact tracing

For COVID-19, the arguments for using cellphone Bluetooth data for contact tracing are

compelling. Because the virus is transmissible through casual contact for at least a few days before onset of symptoms, people are unlikely to be able to recall all those they may have exposed. Even if they could, the number of public health workers needed to perform contact tracing grossly exceeds the available supply. The most likely counterfactual is failure, or inability to fight the epidemic without longer-term social distancing orders. Digital contact tracing offers an effective, less burdensome alternative that is proportional to the threat. Not everyone has smartphones, but using the technology can conserve scarce human resources for working with those who do not.

Should the technology be installed on a mandatory, opt-out, or opt-in basis? The best available evidence suggests that for pathogens with characteristics similar to COVID-19, an opt-out regime is the least restrictive alternative. Research using UK data suggests that COVID-19 could be suppressed if 80% of smartphone users (56% of the overall population, assuming 70% smartphone penetrance) use the app (9). In a survey of U.S. smartphone users, about 40% said they would definitely opt in to a contact tracing app, and just under 70% said definitely or probably (figures were slightly higher for UK users) (10); another survey found much lower opt-in rates among U.S. smartphone users (17% definitely, 32% probably) (11). These numbers fall short of the mark, especially because human inertia will mean some who are willing in theory do not actually install the app. In Singapore—a small nation with fairly high tolerance for government involvement in citizens' lives—only about 20% of smartphone users have installed the TraceTogether app (12).

Thus, although some endorse an opt-in system (3), and legislation proposed in the U.S. would codify that approach (see SM), the available evidence justifies a consent regime that impinges further on individual autonomy. Bolstering the argument in favor of opt out is that users are offered reciprocal benefits: notification if they come into contact with someone dangerous and assistance in protecting friends and family whom they may have endangered.

Although some bioethicists argue that nonconsensual collection of identifiable contact data may be justified for COVID-19 to prevent harm to others (8), opt out is ethically preferable to mandatory use and likely to be sufficient. This setup uses choice architecture to allow those with strong preferences to act on them while not conflating philosophical objections with simple inertia. Studies of electronic health record sharing have found that people tend to stick with the default choice: Only 2 to 5% opt out of health information exchange (13). Although opt outs

from smartphone data sharing may be higher owing to lower trust in government (10), opt out should be tried and evaluated before moving to mandatory use. Opt-out rates can be minimized if public health officials and technology companies collaborate to distribute plain-language FAQs that clearly explain how the data will be collected and used and what benefits there are for users.

Some propose, as an alternative, that contact tracing technology should be mandatory but have a “privacy-protecting” design in which the government receives a list of cases and a list of exposed persons but no information that permits association of particular cases and contacts (14). Such proposals are antithetical to effective epidemiology because they preclude use of the data to track the geographic spread of a pathogen. Use of identifiable data by governments should be carefully limited but must be permitted. TraceTogether operationalizes the least restrictive alternative principle by transmitting users' data to officials only if an individual becomes infected, and then only in specified ways. Executing binding data-use agreements can further ensure data minimization, and experts have articulated several provisions that ought to be included (15). Among these must be the exit strategy—plans for terminating the use of data and destroying data when the public health need for them ends (see SM).

PROCESS RECOMMENDATIONS

When any of these technologies are implemented, it should be through a thoughtful and transparent process. We endorse prior calls for an oversight process by a body that includes members of the public and focuses on particular uses of the data, not just data transfers (3, 8), and for “trustworthy public communication...providing transparent and convincing justifications for the decisions taken” (7) (see SM). South Korea and Taiwan's examples illustrate that diligent transparency can cultivate high levels of trust in the government's strategy (2) (see SM).

But such approaches are far from guaranteed. For example, the situation in the United States leaves much to be desired. A group of technology companies convened by the White House to discuss potential uses of technology to combat COVID-19 has no evident agenda, public or stakeholder group representation, or set of guiding principles. Ethicists and legal experts do not appear to be involved. No processes (for example, adaptation of the notice-and-comment period used for administrative rule-making) have been created for the public to give input. Proposed uses of technology have percolated up through scattered media reports but not through official channels. Communiques from companies working on these technologies are short

on details about how government would be involved. No mechanisms are in place for oversight of tech companies as they pursue this work. The news media and watchdog organizations will continue to be important mechanisms for accountability, but effective oversight requires access to full information about what will be done, how, and why.

Sturdy oversight structures are not easy to stand up in the middle of an emergency. Work will be needed after the COVID-19 threat fades to ensure that we are better prepared next time. For example, federal health agencies could commission a report from the U.S. National Academies of Sciences, Engineering and Medicine recommending rules of the road for digital surveillance in pandemics, and modifications could be made to state and federal privacy and emergency powers laws to facilitate its implementation.

There has been much talk of harnessing the power and ingenuity of the tech sector to fight disease outbreaks, but “harnessing” implies carefully placed constraints and firm direction by a driver. We have yet to craft that yoke. ■

REFERENCES AND NOTES

1. D.J. Hand, *Big Data* 6, 176 (2018).
2. C.J. Wang, C.Y. Ng, R.H. Brook, *JAMA* 10.1001/jama.2020.3151 (2020).
3. L. Ferretti et al., *Science* 368, eabb6936 (2020).
4. Top10VPN, COVID-19 digital rights tracker, 28 April 2020; www.top10vpn.com/research/investigations/covid-19-digital-rights-tracker
5. S. Hendrix, R. Eglash, *Washington Post*, 19 March 2020.
6. E. Vayena et al., *PLoS Comput. Biol.* 11, e1003904 (2015).
7. Nuffield Council on Bioethics, Ethical Considerations in Responding to the COVID-19 Pandemic (Nuffield Council on Bioethics, 2020); www.nuffieldbioethics.org/assets/pdfs/Ethical-considerations-in-responding-to-the-COVID-19-pandemic.pdf
8. Nuffield Council on Bioethics, Guide to the Ethics of Surveillance and Quarantine for Novel Coronavirus (Nuffield Council on Bioethics, 2020); www.nuffieldbioethics.org/assets/pdfs/Guide-to-the-ethics-of-surveillance-and-quarantine-for-novel-coronavirus.pdf
9. R. Hinch et al., Effective configurations of a digital contact tracing app: a report to NHSX (16 April 2020); <https://045.medsci.org.uk/files/files/report-effective-app-configurations.pdf>
10. S. Altmann et al., Acceptability of app-based contact tracing for COVID-19: Cross-country survey evidence (4 May 2020); <https://osf.io/6bn47/?pid=7vgq9>
11. *Washington Post*/University of Maryland, National poll, 21 to 26 April 2020 (5 May 2020); <https://wapo.st/3frvc3v>
12. Z. Doffman, *Forbes* (12 April 2020); <https://bit.ly/35EQjyx>
13. M.M. Mello, J. Adler-Milstein, K.L. Ding, L. Savage, *Milbank Q.* 96, 110 (2018).
14. C. Canca, *Medium*, 10 April 2020; <https://bit.ly/2YJmewn>
15. Z. Emanuel et al., “A national and state plan to end the coronavirus crisis” (Center for American Progress, 2020); www.americanprogress.org/issues/healthcare/news/2020/04/03/482613/national-state-plan-end-coronavirus-crisis

ACKNOWLEDGMENTS

The authors are grateful to C.-M. Chen, I.-M. Parnig, J.-H. Chuang, and H.-W. Jyan for generously sharing information about measures adopted in Taiwan and to G. Lo and G. Wilson for research assistance.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6494/951/suppl/DC1

Published online 11 May 2020

10.1126/science.abb9045



Sentient species, such as elephants, deserve rights, some argue, while less complex creatures do not.

DNA, the rights of parents and children in reproductive scenarios involving genetic engineering, and the challenges presented by animal-human chimeras. Science-literate readers might find this chapter a bit underwhelming, as it is written to be accessible to a broad audience.

The book's fifth chapter suggests that rights could be employed to protect vulnerable groups suffering under corrupt leadership, which, while generally illegal, is frequently tolerated, often at the expense of the poor and powerless. Notably, the authors argue for a right to be free from corruption rather than the right to good governance, as they see corruption as much broader than an abuse of authority. They include, for example, financial rules that intensify the unfair distribution of wealth.

A recurring theme in the book is that of competing rights. For example, some advocates have proposed that same-sex couples should be guaranteed the right to access surrogacy services, but such a right could challenge the rights and autonomy of surrogate mothers, an often-oppressed group who are potential victims of trafficking and exploitation. Schulz and Raman also describe antagonistic interactions between competing rights groups, for example, between some women's rights advocates and groups promoting the rights of nonbinary individuals. The authors suggest that "the inclusion of newly established rights need not vitiate the old."

The book closes with a discussion of the possible extension of human rights to nonhumans, including animals, robots, and nature. These chapters provide an excellent

overview of current discussions but offer little in the way of novel and provocative rights. Here also, the authors are forced to make distinctions between those animals, robots, and natural resources that ought to be granted rights and those that should not, owing to conflicting needs and demands. Such distinctions are sometimes rational, as in the case of the rights of obligate carnivores over their prey, and at other times seemingly arbitrary, as when Schulz and Raman sug-

gest valuing the survival of native species over invasive ones that were introduced to a habitat through no fault of their own. One is left wondering whether these proposed divisions will one day seem as problematic and discriminatory as others the authors argue against elsewhere in the book. ■

10.1126/science.abc0163



The Coming Good Society
William F. Schulz
and Sushma Raman
Harvard University Press,
2020. 328 pp.

oppression, and provide a sense of dignity to groups within these penumbras.

In the book's third chapter, Schulz and Raman consider new rights to which they believe we are all entitled, in the form of improved privacy protections. They ground these rights in ideals of autonomy and control, as corporations seek to learn more about us and to use that information to manipulate our behavior. The authors propose the widespread implementation of rights already granted—recently and prominently—to Europeans through the General Data Protection Regulation. These include the right to be forgotten (i.e., the right to have negative personal data erased from internet search results). Schulz and Raman also advocate using algorithm transparency requirements to limit the biases that are often unintentionally encoded into those algorithms, which reflect the "presumptions and predictions of their makers."

The authors continue their discussion of emerging privacy concerns in the book's fourth chapter, which examines how privacy should be accorded to our genetic information. Here, they address the ownership of

BOOKS *et al.*

POLICY

Revisiting rights in an ever-evolving world

Data and genetic privacy join proposed privileges for nonhumans, robots, and nature

By **Dov Greenbaum**

The *Coming Good Society* offers a cursory overview of the state of human rights in an age of emerging technologies. More accessible narrative than academic treatise, this enjoyable read examines how changing norms create opportunities to expand the scope of universal protections and rights. Authors William F. Schulz, former executive director of Amnesty International USA, and Sushma Raman, executive director of the Harvard Kennedy School Carr Center for Human Rights Policy, look to their own experiences and extensive professional encounters to support their arguments in the book's eight short but substantive chapters.

Rights reflect our consensus of what is or ought to be in a good society. As such, human rights are in constant flux, evolving and sometimes devolving as social and political norms change, societies adapt, and discoveries and advancing technologies change worldviews. Rights are also transactional claims that create duties and obligations of the powerful to the less powerful. They promote civility, deter

The reviewer is at Zvi Meitar Institute for Legal Implications of Emerging Technologies, Herzliya, Israel, and the Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT, USA. Email: dov.greenbaum@idc.ac.il

PHOTO: ISTOCK.COM/EM CAMPOS

CLIMATE POLICY

The will to act when the data are dire

Then and now, American politicians have failed to heed scientific warnings

By Erika Lorraine Milam

In the face of a dire planetary prognosis, American politicians in the 1980s faced a simple question: Act now, or wait and see? These are the same options available to anyone presented with an uncertain future, whether rising levels of atmospheric carbon dioxide, a cyst that might turn cancerous, or scattered cases of an infectious disease that could explode into a pandemic. *Losing Earth* focuses on the first of these, but as I read the book—under a stay-at-home order in New Jersey owing to the spread of coronavirus disease 2019—Nathaniel Rich's frustration resonated with my own. Why, he asks, did American politicians fail to act when scientists agreed that there was a real and present danger?

In 1951, environmentalist Rachel Carson had noted that the growing season in the subarctic regions was longer, growth rings on trees were fatter, and cod were migrating farther north. "The long trend is toward a warmer earth," she wrote (1).

In 1975, anthropologist Margaret Mead convened a symposium calling attention to the endangered atmosphere. Arguing that the issue would need to be addressed on a planetary scale, she called for more research, especially scientific models of the likely future that could guide action in the present.

Rich observes that by 1979, scientists had assembled the essential pieces of the climate warming puzzle. By the end of the next decade, evidence regarding the future of the planet was incontrovertible: Increased levels of carbon dioxide in the atmosphere were causing the planet to warm. Models warned of a litany of dangers: rising sea levels, the abandonment of coastal cities, widespread droughts, and shifting agricultural belts. These issues, Rich notes, were understood to be environmental in nature and political in impact.

The book's two main characters—Rafe Pomerance and James Hansen—move through its pages like characters from a Frederick Forsyth novel: cool, competent, unfussy, and at odds with a world that underappreciates their hard-won expertise. Pomerance was "not a scientist" but a stooping six-foot-

four environmental policy analyst with horn-rimmed glasses and a mustache. Hansen had majored in math and physics but dreamed of baseball. He helped create computer models of carbon circulation on a planetary scale and worked with Jule Charney to produce a report that predicted that the planet would warm by 3°C in the next century (2). Pomerance scrutinized the evidence, read between the lines, and started calling politicians.

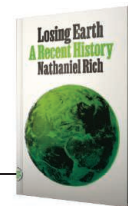
They might have succeeded in convincing those in power of the need to curb the world's fossil fuel dependency were it not for President Ronald Reagan's investment in fossil fuels and unrelenting war on environmentalism throughout the 1980s, followed by President George H. W. Bush's 1989 appointment of John Sununu as White House Chief of staff. If Pomerance and Hansen are the book's heroes, Sununu is the archvillain.

Ever since Mead's 1975 conference, Rich posits, Sununu had interpreted claims of global warming as an excuse to bridle economic progress and enact authoritarian global solutions to a problem whose existence he doubted. Pomerance and Hansen could not convince Sununu otherwise, and when given the opportunity to steer the country's policy away from reckoning with climate change, Sununu took it. (He remains skeptical of global warming to this day.)

Rich's tight focus on the lives of a handful of men allows him to frame science and politics as mirror worlds—Hansen and Pomerance never quite grokking the power dynamics

Losing Earth: A Recent History

Nathaniel Rich
Farrar, Straus and Giroux,
2019. 224 pp.



of politics and Sununu handily rejecting the logic of science. More broadly, Rich suggests that many Americans found it difficult to appreciate the connection between actions in the present and long-term effects, the lag between cause and effect exacerbated by our species's tolerance for self-delusion.

Rich's description of the politicians' inaction when presented with robust scientific evidence now feels prescient. Yet from the perspective of a world beset by the social, political, and health impacts of a novel coronavirus, I found myself searching the book for the key relations that characterize the causal space between personal choices and a cultural zeitgeist. It was not until the epilogue that Rich turned his attention to the powerful dynamics of funding in science, attempts of the fossil fuel industry to manipulate the optics of knowledge for a broader public, and the disproportionate effects of climate change according to demographics and geography. ■

REFERENCES AND NOTES

1. R. Carson, *The Sea Around Us* (Oxford Univ. Press, 1951).
2. National Research Council, *Carbon Dioxide and Climate: A Scientific Assessment* (The National Academies Press, 1979); <https://doi.org/10.17226/12181>.

10.1126/science.abc6171



Fertility rates and other indicators of human progress have declined.

PODCAST

Slowdown: The End of the Great Acceleration—and Why It's Good for the Planet, the Economy, and Our Lives

Danny Dorling
Yale University Press, 2020. 400 pp.

Even before the pandemic brought many institutions to an abrupt standstill, a number of indicators—from fertility rates to GDP growth—suggested that human progress has slowed in recent decades. This week on the *Science* podcast, geographer Danny Dorling reveals why this might not be such a bad thing.

sciencemag.org/podcasts

10.1126/science.abc6198

The reviewer is at the Department of History, Princeton University, Princeton, NJ 08544, USA. Email: emilam@princeton.edu



LETTERS

Unnecessary restrictions on blood donors should be removed to maximize the blood and plasma available for use.

Edited by Jennifer Sills

Ease restrictions on U.S. blood donations

With a vaccine for coronavirus disease 2019 (COVID-19) likely more than a year away, we must identify effective therapies for patients now. One promising approach is the use of plasma from patients who have recovered from COVID-19 (1, 2). To facilitate this strategy, the U.S. Food and Drug Administration (FDA) recently revised some of the restrictions on blood donation, including a decrease in deferral time for men who have sex with men (MSM) to 3 months (3). This is a positive change to an outdated guideline, but it does not go far enough.

In 1983, the FDA indefinitely barred all MSM from donating blood for fear of transmitting human immunodeficiency virus (HIV) and hepatitis B/C by transfusion. In 2015, the lifetime ban was changed to 12 months from last sexual contact (4). Today, the risk of contracting HIV or hepatitis B/C through transfusion is less than 1 in 2 million, and the incidence is substantially lower (5). This success is due to advances in screening, not to banning MSM from donating blood. The false-negative rates of modern HIV nucleic acid tests fall around 0.05%. The window between infection and detection has dropped to 9 days (5, 6). Despite this improvement, the FDA continues to exclude otherwise healthy MSM through arguably discriminatory policy. Although a step forward from the 12-month policy, a deferral period of 3 full months is not necessary to protect patients (7, 8).

Moreover, this revised policy may not meaningfully increase the donor pool, given that waiting until 3 months after sexual contact amounts to a lifetime blood donation ban for many men.

The demand for healthy blood and convalescent plasma will accelerate as COVID-19 infects more Americans. To address the acute shortage (9), the deferral period should be decreased to 2 weeks, after which we can reliably screen for HIV. More granular deferrals could also be introduced. Instead of a blanket discriminatory ban on MSM blood donations, we could evaluate donors based on concrete risky behaviors, such as having unprotected sex with multiple sexual partners or sharing needles. Alternatively, we could inactivate pathogenic DNA and RNA with intercalating molecules such as amotosalen, which European blood centers already do routinely (10, 11). Coupled with robust testing and screening, these approaches will exclude fewer healthy donors while still minimizing the risk of transfusion-transmitted HIV.

The FDA's policies must be grounded in science. Safely lifting the restrictions on blood donations has the potential to save millions of lives in a normal year (12). Now, plasma may play a crucial role in treating patients suffering from COVID-19. We cannot afford to turn away HIV-negative blood with lifesaving antibodies just because the donor is gay or bisexual.

Chirag Vasavda^{1*}, Byron K. Ho², Solomon H. Snyder¹

¹The Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. ²Department of Dermatology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.

*Corresponding author. Email: cvasavda@jhmi.edu

REFERENCES AND NOTES

1. A. Casadevall, L.-A. Pirofski, *J. Clin. Invest.* **130**, 1545 (2020).
2. C. Shen *et al.*, *JAMA* **323**, 1582 (2020).
3. P. Marks, "Coronavirus (COVID-19) update: FDA provides updated guidance to address the urgent need for blood during the pandemic" (FDA, 2020); www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-provides-updated-guidance-address-urgent-need-blood-during-pandemic.
4. FDA, "Revised recommendations for reducing the risk of Human Immunodeficiency Virus transmission by blood and blood products" (2020); www.fda.gov/regulatory-information/search-fda-guidance-documents/revised-recommendations-reducing-risk-human-immunodeficiency-virus-transmission-blood-and-blood.
5. S. Zou, S. L. Stramer, R. Y. Dodd, *Transfus. Med. Rev.* **26**, 119 (2012).
6. M. Busch, K. Watanabe, J. Smith, S. Hermansen, R. Thomson, *Transfusion* **40**, 585 (2000).
7. E. DeBuck, T. Deltjens, V. Compennolle, P. Vandekerckhove, *PLOS One* **10**, e0122523 (2015).
8. B. R. Sturrock, S. Mucklow, *Clin. Med.* **18**, 304 (2018).
9. J. Stephenson, "Canceled blood drives, social distancing cause nationwide blood shortages," *JAMA Health Forum* (2020).
10. L. Lin *et al.*, *Transfusion* **45**, 580 (2005).
11. J. Irsch, L. Lin, *Transfus. Med. Hemother.* **38**, 19 (2011).
12. A. J. Miyashita, G. J. Gates, *Effects of Lifting Blood Donation Bans on Men Who Have Sex with Men* (Williams Institute, 2014).

10.1126/science.abc2322

The precarious position of postdocs during COVID-19

Postdoctoral researchers play a crucial role in many research groups, serving as mentors, teachers, and leaders as they develop their skills and prepare for scientific careers (1). However, the coronavirus disease 2019 (COVID-19) crisis has put funding and support for postdoc positions at risk, threatening to upend the career paths available to these junior scientists.

Even in normal times, postdoctoral positions provide little job security (2, 3). Most postdocs are employed on yearly contracts, and the availability of research funds is highly variable (1). Postdocs receive little institutional support in comparison to undergraduates, graduate students, and faculty. The positions often do not provide access to affordable health care or child care, career counseling or resources, paid sick leave, or employee and student benefits such as alumni network membership or union representation (1, 2, 4–7). Before the pandemic, an ongoing national discussion among postdocs was taking place to address the benefits of collective bargaining and unionization, as many feel the working conditions and terms of employment are substandard or outright nonexistent (1, 3, 8). For 2 to 3 years (and sometimes much longer), postdocs tolerate these subpar conditions in hopes of using their experience to propel them into full-time jobs as professors or researchers outside academia (2).

However, the economic crisis resulting from COVID-19 stay-at-home orders has spurred a growing list of universities to implement hiring freezes and cancel new faculty hires (9, 10). This lowers the chances that postdocs can obtain coveted full-time positions. Meanwhile, experimental work has all but ground to a halt, visas are expiring with little clarity about the prospect of extensions, and continued funding has become uncertain, jeopardizing the time-sensitive research that postdocs are conducting during their short contracts.

Although many institutions have granted some form of pandemic relief to other members of the academic community, postdocs have been overlooked (11, 12). To protect the future and diversity of the scientific pipeline, universities and research institutes must take immediate action to retain these vital junior scientists. Institutions should implement programs to prolong fellowship positions (12), similar to stop-the-clock policies available to tenure-track faculty, and vigorously advocate for federal-level extensions to visa programs. They should also offer temporary assistance to help vulnerable postdocs cope with current child care and health care challenges.

M. Arslan Ahmed, Amir H. Behbahani*, Adrian Brückner, Caroline J. Charpentier, Livia H. Morais, Stewart Mallory, Allan-Hermann Pool
California Institute of Technology, Pasadena, CA 91125, USA.

*Corresponding author. Email: amirhb@caltech.edu

REFERENCES AND NOTES

1. S. C. McConnell *et al.*, *eLife* **7**, e40189 (2018).
2. S. Kahl, D. K. Ginther, *Nat. Biotechnol.* **35**, 90 (2017).
3. H. Sauerermann, M. Roach, *Science* **352**, 663 (2016).
4. M. Alun *et al.*, *Nat. Ecol. Evol.* **10**, 1038/s41559-020-1178-6 (2020).

5. C. Woolston, *Nature* **565**, 125 (2019).
6. K. Spratt, "Affordable child care for postdocs: One institution's solution," *Science* (2000).
7. M. Kristoffersen, "Postdocs struggle with child care costs," *Yale News* (2020).
8. B. D. Weitzner, "Why I'm pushing for a postdoc union," *Science* (2017).
9. K. Kelsky, "The professor is in: Stranded on the academic job market this year?," *The Chronicle of Higher Education* (2020).
10. M. T. Nietzel, "College furloughs have begun," *Forbes* (2020).
11. N. Gott, "UC Berkeley students receive emergency aid as part of federal coronavirus relief bill," *The Daily Californian* (2020).
12. Ed Markey, United States Senator for Massachusetts, Press Release, "Senators Markey and Tillis lead 31 senators in calling for \$26 billion in next coronavirus relief package to support research workforce" (2020).

COMPETING INTERESTS

All authors are board members of the California Institute of Technology (Caltech) Postdoctoral Association (CPA). The views expressed herein do not in any way represent the view of other members of the CPA board, postdoctoral individuals at Caltech, or Caltech.

10.1126/science.abc5143

Sumatran rhinoceros on the brink of extinction

The Sumatran rhinoceros (*Dicerorhinus sumatrensis*), the closest living relative to the extinct woolly rhino, has been on decline for about a million years (1), but it is now at risk of imminent extinction. According to the International Union for Conservation of Nature Red List, the Sumatran rhino is Critically Endangered (2). In just 20 years, the species population has decreased from 250 to just 80 animals. Since the recent death of the last Sumatran rhino in Malaysia (3), all remaining individuals live in one of four subpopulations in Indonesia (2, 4). The current population is not sustainable without the help of breeding programs.

The main reason for the Sumatran rhino's population collapse is poaching driven by the Asian black market of traditional medicine, where a kilo of rhino horn can sell at around US\$65,000 (5, 6). In addition, human activities such as deforestation fragment the rhino's habitats (4). If these human activities continue, the rhino population will likely go extinct by 2030 (7).

In addition to curtailing these harmful activities, capture, relocation, and breeding programs are now critical to prevent population collapse and avoid harmful mutations leading to diseases that reduce the reproductive capacity of the Sumatran rhinos (8, 9). The breeding programs managed by the World Association of Zoos and Aquariums (10) may be able to provide conservation programs with new individuals to increase genetic diversity. However, these strategies must take into account that Sumatran rhinos do not thrive or breed well in captivity or outside their ecosystem (11). Breeding

programs should urgently be established in the rhino's natural habitat and include both natural and artificial insemination as well as embryo technologies, as has been tried for the northern white rhino (12). We must devote time and resources to ensure that the remaining 80 Sumatran rhinos are not the last.

Su Shiung Lam^{1,2}, Nyuk Ling Ma^{2,1}, Wanxi Peng¹, Christian Sonne^{3,1*}

¹Henan Agricultural University, Zhengzhou, China.

²Universiti Malaysia Terengganu, Terengganu, Malaysia.

³Aarhus University, Roskilde, Denmark.

*Corresponding author. Email: cs@bios.au.dk

REFERENCES AND NOTES

1. H. L. Mays *et al.*, *Curr. Biol.* **28**, 70 (2018).
2. N. J. van Strien *et al.*, *Dicerorhinus sumatrensis* (The IUCN Red List of Threatened Species, 2008).
3. J. Bittel, "Last Sumatran rhino in Malaysia dies," *National Geographic* (2019).
4. International Rhino Foundation, Sumatran rhino (<https://rhinos.org/species/sumatran-rhino/>).
5. H. H. T. Prins, B. Okita-Ouma, *Science* **340**, 1167 (2013).
6. L. Neme, "Despite ban, rhino horn flooding black markets across China," *National Geographic* (2017).
7. W. J. Ripple *et al.*, *Sci. Adv.* **1**, e1400103 (2015).
8. International Rhino Foundation, Sumatran rhino rescue (<https://rhinos.org/sumatranrhinorecue/>).
9. N. E. Schaffer *et al.*, *J. Threat. Taxa* **12**, 15279 (2020).
10. World Association of Zoos and Aquariums, Conservation Breeding Programmes (www.waza.org/priorities/conservation-breeding-programmes/).
11. J. Hance, "When it comes to captive breeding, not all Sumatran rhinos are equal," *Mongabay* (2019).
12. E. Callaway, *Nature* **533**, 20 (2016).

10.1126/science.abc2202

TECHNICAL COMMENT ABSTRACTS

Comment on "High-surface-area corundum by mechanochemically induced phase transformation of boehmite"

Jiangong Li, Sanxu Pu, Wenbin Cao, Lu Li, Ruiyun Guo

Amrute *et al.* (Reports, 25 October 2019, p. 485) claimed that no methods were able to produce high-purity α -Al₂O₃ with surface areas greater than 100 m² g⁻¹, even though much higher surface areas up to 253 m² g⁻¹ have been reported. Moreover, the materials they obtained could be porous aggregates and may not be 13-nm nanoparticles, as claimed.

Full text: [dx.doi.org/10.1126/science.abb0142](https://doi.org/10.1126/science.abb0142)

Response to Comment on "High-surface-area corundum by mechanochemically induced phase transformation of boehmite"

Amol P. Amrute, Zbigniew Łodziana, Hannah Schreyer, Claudia Weidenhauer, Ferdi Schütt

Li *et al.* commented that our report claims that methods reported thus far cannot enable the production of high-purity corundum with surface areas greater than 100 m² g⁻¹, and that our obtained material could be porous aggregates rather than nanoparticles. We disagree with both of these suggestions.

Full text: [dx.doi.org/10.1126/science.abb0948](https://doi.org/10.1126/science.abb0948)

Cite as: J. Li *et al.*, *Science*
10.1126/science.abb0142 (2020).

Comment on “High-surface-area corundum by mechanochemically induced phase transformation of boehmite”

Jiangong Li*, Sanxu Pu, Wenbin Cao, Lu Li, Ruiyun Guo

Institute of Materials Science and Engineering, Lanzhou University, Lanzhou 730000, China.

*Corresponding author. Email: lijg@lzu.edu.cn

Amrute *et al.* (Reports, 25 October 2019, p. 485) claimed that no methods were able to produce high-purity α -Al₂O₃ with surface areas greater than 100 m² g⁻¹, even though much higher surface areas up to 253 m² g⁻¹ have been reported. Moreover, the materials they obtained could be porous aggregates and may not be 13-nm nanoparticles, as claimed.

Recently, Amrute *et al.* (1) reported the synthesis of α -Al₂O₃ nanoparticles (diameter ~13 nm, surface areas ~140 m² g⁻¹) by the mechanochemical dehydration of hydrothermally synthesized boehmite at room temperature. They claimed that no methods were available for the fabrication of high-purity α -Al₂O₃ with surface areas exceeding 100 m² g⁻¹. However, higher surface areas between 161 and 253 m² g⁻¹ have been reported for α -Al₂O₃ nanoparticles synthesized by several methods (2–4). The first example (2) synthesized α -Al₂O₃ nanoparticles embedded in Fe matrix from coarse Al and Fe₂O₃ powders by a mechanochemical method, followed by a HCl treatment to obtain disperse single-phase α -Al₂O₃ with 99.6% purity by weight and an average particle size of 14.3 nm. α -Al₂O₃ nanoparticles with ultrafine sizes and narrow size distributions could easily be separated by fractionated coagulation using HCl as coagulating agent. Average particle sizes of 5.2, 6.5, 7.9, and 9.6 nm were obtained. The surface area of the 7.9-nm α -Al₂O₃ nanoparticles was 170 m² g⁻¹, measured by N₂ adsorption at 77 K using the Brunauer-Emmett-Teller method. The second example (3) synthesized a mixture of α -Al₂O₃ and α -Fe₂O₃ with 1:5 molar ratio by coprecipitation and calcination, followed by a HCl corrosion to obtain disperse single-phase α -Al₂O₃ with 98.28% purity by weight (main impurity: Fe), an average particle size of 9 nm, and a surface area of 161 m² g⁻¹. The third example (4) synthesized α -Al₂O₃ nanoparticles by direct ball milling of micrometer-sized α -Al₂O₃ powders, followed by a HCl treatment to obtain disperse single-phase α -Al₂O₃ with 99.96% purity by weight and an average particle size of 8 nm. α -Al₂O₃ nanoparticles with ultrafine sizes and narrow size distributions could easily be fractionated using HCl as

coagulating agent. Average particle sizes of 3.3 nm (Fig. 1) (4), 4.8 nm, and 6.8 nm, respectively, were obtained. The surface area of the 4.8-nm α -Al₂O₃ nanoparticles was 253 m² g⁻¹. All three methods produced high-purity, dispersed non-aggregated α -Al₂O₃ nanoparticles (Fig. 1) with surface areas higher than that reported by Amrute *et al.*

Amrute *et al.* claimed that the synthesized α -Al₂O₃ consisted of nanoparticles with an average diameter of ~13 nm. However, no individual nanoparticles with a diameter of ~13 nm can be identified in their transmission electron microscopy (TEM) images [figure 4, A and C, of (1)] and their α -Al₂O₃ could be porous aggregates. According to the standard definitions of particles, agglomerates (weakly bonded by van der Waals forces), and aggregates (bonded by strong chemical bonds) (5), what can be observed in the TEM images [figure 4, A and C, of (1)] are ~200-nm porous aggregates consisting of ~13-nm primary particles and nanopores. They should not be softly agglomerated together, because agglomerates can be separated into individual nanoparticles by the electron beam in TEM analysis.

The high surface areas of the α -Al₂O₃ reported in (1) may result from the nanostructure and special pore channels of the mechanochemically dehydrated product of boehmite. The decomposition of boehmite induced by ball milling may yield a nanocomposite of ~41 vol % liquid water and ~59 vol % α -Al₂O₃. During calcination at 823 K, after evaporation of water, the remaining α -Al₂O₃ could become porous aggregates with ~41 vol % mesopores. The pore size distribution and the type IV N₂ adsorption-desorption isotherms with an H1-type hysteresis loop [according to the International Union of Pure and Applied Chemistry (6)] in

figure 4E of (I) are indicative of open-ended cylindrical mesopores. The dynamic light-scattering measurements [figure 2B of (I)] reveal that most α -Al₂O₃ aggregates are ~30 nm in size. Hence, we believe that the high-surface-area α -Al₂O₃ reported in (I) may be 30- to 200-nm mesoporous aggregates, instead of ~13-nm nanoparticles as claimed by Amrute *et al.*

REFERENCES

1. A. P. Amrute, Z. Łodziana, H. Schreyer, C. Weidenthaler, F. Schüth, High-surface-area corundum by mechanochemically induced phase transformation of boehmite. *Science* **366**, 485–489 (2019). [doi:10.1126/science.aaw9377](https://doi.org/10.1126/science.aaw9377) [Medline](#)
2. S. Pu, L. Li, J. Ma, F. Lu, J. Li, Disperse fine equiaxed alpha alumina nanoparticles with narrow size distribution synthesised by selective corrosion and coagulation separation. *Sci. Rep.* **5**, 11575 (2015). [doi:10.1038/srep11575](https://doi.org/10.1038/srep11575) [Medline](#)
3. R. Guo, W. Cao, X. Mao, J. Li, Selective corrosion preparation and sintering of disperse α -Al₂O₃ nanoparticles. *J. Am. Ceram. Soc.* **99**, 3556–3560 (2016). [doi:10.1111/jace.14403](https://doi.org/10.1111/jace.14403)
4. L. Li, S. Pu, Y. Liu, L. Zhao, J. Ma, J. Li, High-purity disperse α -Al₂O₃ nanoparticles synthesized by high-energy ball milling. *Adv. Powder Technol.* **29**, 2194–2203 (2018). [doi:10.1016/j.appt.2018.06.003](https://doi.org/10.1016/j.appt.2018.06.003)
5. A. Balakrishnan, P. Pizette, C. L. Martin, S. V. Joshi, B. P. Saha, Effect of particle size in aggregated and agglomerated ceramic powders. *Acta Mater.* **58**, 802–812 (2010). [doi:10.1016/j.actamat.2009.09.058](https://doi.org/10.1016/j.actamat.2009.09.058)
6. K. S. W. Sing, D. H. Everett, R. A. W. Haul, L. Moscou, R. A. Pierotti, J. Rouquérol, T. Siemienińska, Reporting physisorption data for gas/solid systems with special reference to the determination of surface area and porosity (Recommendations 1984). *Pure Appl. Chem.* **57**, 603–619 (1985). [doi:10.1351/pac198557040603](https://doi.org/10.1351/pac198557040603)

ACKNOWLEDGMENTS

We thank Y. Dong (MIT) for stimulating discussion. **Funding:** Supported by the National Natural Science Foundation of China (51272098, 51551201, and 51772137) and the Fundamental Research Funds for the Central Universities (Izujbky-2019-sp03). **Competing interests:** The authors have filed patent applications on the synthesis and size-selective separation processes of disperse fine equiaxed α -alumina nanoparticles with narrow size distributions.

23 January 2020; accepted 6 May 2020
Published online 29 May 2020
[10.1126/science.abb0142](https://doi.org/10.1126/science.abb0142)

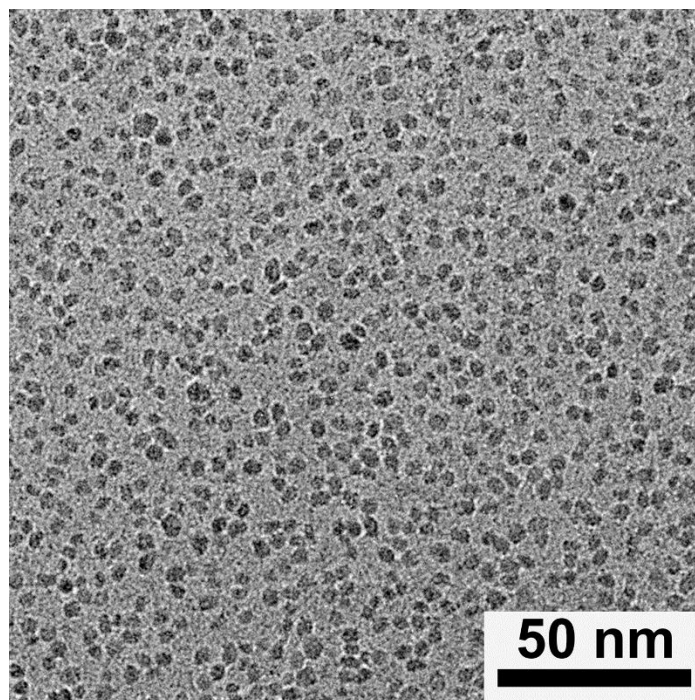


Fig. 1. Disperse fine α - Al_2O_3 nanoparticles with a narrow size distribution. Shown is a TEM image of α - Al_2O_3 nanoparticles with an average size of 3.3 nm and a size distribution of 2 to 6 nm. [Reproduced, with permission, from figure S3A of (4)]

Cite as: A. P. Amrute *et al.*, *Science*
10.1126/science.abb0948 (2020).

Response to Comment on “High-surface-area corundum by mechanochemically induced phase transformation of boehmite”

Amol P. Amrute^{1*}, Zbigniew Łodziana², Hannah Schreyer¹, Claudia Weidenthaler¹, Ferdi Schüth^{1*}

¹Department of Heterogeneous Catalysis, Max-Planck-Institut für Kohlenforschung, D-45470 Mülheim an der Ruhr, Germany. ²INP, Polish Academy of Sciences, PL-31-342 Kraków, Poland.

*Corresponding author. Email: amrute@kofo.mpg.de (A.P.A.); schueth@kofo.mpg.de (F.S.)

Li *et al.* commented that our report claims that methods reported thus far cannot enable the production of high-purity corundum with surface areas greater than 100 m² g⁻¹, and that our obtained material could be porous aggregates rather than nanoparticles. We disagree with both of these suggestions.

High-surface-area corundum (α -Al₂O₃) with nanometer-sized particles is a highly desirable material from both industrial and academic viewpoints (1, 2). However, the synthesis of high-purity α -Al₂O₃ in nanoparticulate form is not so straightforward because of the high activation barrier for its formation from transition aluminas. Previously studied methods reporting the production of high-surface-area corundum suffer from the requirements of harsh conditions either in the pre-synthesis step (2–4) (i.e., to obtain an appropriate precursor, such as diaspora, which can be converted to corundum at lower temperature) or in the post-synthesis step to remove the seed materials, matrix, or contamination (e.g., iron or iron oxide) in which corundum is formed. When micrometer-sized preformed corundum is milled, it is also severely contaminated with milling media and jar material. Our approach reported in (1) presents a simple procedure, enabling easy access to high-surface-area α -Al₂O₃, from which impurities, if any, can easily be removed, or can be avoided altogether, if milling proceeds in corundum milling equipment. Our report in (1) (i) describes ball milling of boehmite (γ -AlOOH) at room temperature to produce nano-corundum, and (ii) sheds light on the microstructural changes, and on the chemistry and thermodynamic aspects, involved in the transformation of γ -AlOOH to α -Al₂O₃.

We do not agree with the points raised by Li *et al.* (5) concerning our paper (1), although we should have included citations to their recent papers on the synthesis of α -Al₂O₃ nanoparticles (NPs) as prior art (6–8). The omission was unintentional and has been corrected. Their work did not appear in our searches, likely owing to the different scope of our study. Also, the conversion of transition aluminas by

phase transformation, or of preformed large-crystal conventional α -Al₂O₃ by top-down ball-milling approaches, had been described somewhat more extensively before (9–12).

Regarding the first point raised by Li *et al.*, we have not made any strong claim that there is no method available to produce α -Al₂O₃ with high surface area (>100 m² g⁻¹). Instead, we have pointed out the difficulties of achieving high-surface-area α -Al₂O₃. In fact, we have included an example of the diaspora route, which allows the production of corundum at mild calcination temperatures and thus enables the preservation of high surface area. This route has long been known to produce α -Al₂O₃ with surface areas exceeding 150 m² g⁻¹ (2). However, the production of precursor diaspora is a difficult process (3, 4), which makes this route technically unattractive.

Regarding the second point raised by Li *et al.* on the size of ball milling-derived corundum NPs, we have additional evidence, in addition to transmission electron microscopy (TEM), that supports the particle sizes we have reported. The TEM images may suffer from the sample preparation, deposition on the sample holder, or both (13, 14). As we had mentioned, our sample preparation includes a sprinkling of dry sample specimens on the TEM grid. We chose this approach because we wanted to analyze the sample directly after milling or after mild calcination of the ball-milled sample. Such sample preparation could lead to the deposition of an excessive amount in a less dispersed manner. Therefore, what appears in the TEM images as agglomerates are the overlapped particles; they are certainly not aggregates of ~200 nm.

To substantiate this, we have repeated our TEM analysis by varying the sample preparation. In this case, we used

ethanol as a solvent to disperse the sample and ultrasonicated it for ~20 min, then placed a drop (2 μ l) on the TEM grid and allowed it to dry under ambient conditions. Such sample preparation is also routinely used for microscopy analysis. The obtained images visualize well-separated NPs in the range of 3 to 30 nm (Fig. 1). The NPs might be separated even more if we were to reduce the concentration of the sample in ethanol during the sample preparation and selectively take a sample from the solution phase during ultrasonication. However, this procedure might not lead to a realistic picture, as a much smaller fraction of the sample will be visualized. We emphasize that we analyzed the materials directly after ball milling (or after mild calcination of the ball-milled sample) and did not process it further by methods such as fractional coagulation, which discards larger NPs and selectively collects the smaller ones from the acid solution phase (6, 7).

In addition to TEM, dynamic light scattering (DLS) analysis shows hydrodynamic diameter after 3-hour milling of ~26 nm [see figure 2B of (1), blue curve, data point for 180-min milled sample]. Therefore, if one excludes the contribution of the charged layer in this value, our particle diameter will fall in the range we observed from TEM images. DLS is a bulk method, providing a more comprehensive picture of particle size. Moreover, size analysis from x-ray diffraction (XRD) also provides an average domain size of ~18 nm [table 1 of (1)]. Thus, we are convinced that our sample is mostly composed of corundum NPs of ~15 nm (average particle diameter), in agreement with the value for surface area.

Regarding porosity, it is highly unlikely that we have intraparticle pores or channels. If such voids were formed, they would collapse under the high compression force applied in the high-intensity ball-milling process. There is also no indication of pores or channels from the different analytical methods. Still, we cannot exclude any contribution from surface roughness to the observed surface area values. Some surface roughness could indeed be seen in the TEM images shown in (1). We hope that this explanation further clarifies that our sample is mostly composed of small NPs that give the high surface area. Note that particle sizes obtained from TEM analysis, sorption, DLS, and XRD reflection broadening agree quite well, which supports the statement made above.

We want to emphasize that the simplicity of our process provides greater prospects for scalability. In fact, we have already achieved success in scaling up to several hundred grams in a mill type called Simoloyer. Because Simoloyer is available with milling jar sizes ranging from 1 to 900 liters, it provides good perspectives for further scale-up. In contrast, scale-up of the process reported by Li and colleagues (6) might not be straightforward or cost-effective, consider-

ing their starting materials, milling time, and harsh post-synthesis separation process.

REFERENCES

1. A. P. Amrute, Z. Łodziana, H. Schreyer, C. Weidenthaler, F. Schüth, High-surface-area corundum by mechanochemically induced phase transformation of boehmite. *Science* **366**, 485–489 (2019). [doi:10.1126/science.aaw9377](https://doi.org/10.1126/science.aaw9377) [Medline](#)
2. J. M. McHale, A. Auroux, A. J. Perrotta, A. Navrotsky, Surface energies and thermodynamic phase stability in nanocrystalline aluminas. *Science* **277**, 788–791 (1997). [doi:10.1126/science.277.5327.788](https://doi.org/10.1126/science.277.5327.788)
3. H. Yanagida, G. Yamaguchi, A discussion on the phase diagram of the system $\text{Al}_2\text{O}_3\text{-H}_2\text{O}$ considering the transformation mechanism of the polymorphs appearing in it. *J. Ceram. Assoc. Jpn.* **74**, 94–100 (1966). [doi:10.2109/jcersj1950.74.847_94](https://doi.org/10.2109/jcersj1950.74.847_94)
4. T. Tsuchida, K. Kodaira, Hydrothermal synthesis and characterization of diaspor, $\beta\text{-Al}_2\text{O}_3\text{-H}_2\text{O}$. *J. Mater. Sci.* **25**, 4423–4426 (1990). [doi:10.1007/BF00581103](https://doi.org/10.1007/BF00581103)
5. J. Li, S. Pu, W. Cao, L. Li, R. Guo, Comment on “High-surface-area corundum by mechanochemically induced phase transformation of boehmite”. *Science* **368**, eabb0142 (2020).
6. S. Pu, L. Li, J. Ma, F. Lu, J. Li, Disperse fine equiaxed alpha alumina nanoparticles with narrow size distribution synthesised by selective corrosion and coagulation separation. *Sci. Rep.* **5**, 11575 (2015). [doi:10.1038/srep11575](https://doi.org/10.1038/srep11575) [Medline](#)
7. R. Guo, W. Cao, X. Mao, J. Li, Selective corrosion preparation and sintering of disperse $\alpha\text{-Al}_2\text{O}_3$ nanoparticles. *J. Am. Ceram. Soc.* **99**, 3556–3560 (2016). [doi:10.1111/jace.14403](https://doi.org/10.1111/jace.14403)
8. L. Li, S. Pu, Y. Liu, L. Zhao, J. Ma, J. Li, High-purity disperse $\alpha\text{-Al}_2\text{O}_3$ nanoparticles synthesized by high-energy ball milling. *Adv. Powder Technol.* **29**, 2194–2203 (2018). [doi:10.1016/j.apt.2018.06.003](https://doi.org/10.1016/j.apt.2018.06.003)
9. P. A. Zieliński, R. Schulz, S. Kaliaguine, A. Van Neste, Structural transformations of alumina by high energy ball milling. *J. Mater. Res.* **8**, 29852992 (1993).
10. S. R. Chauruka, A. Hassanpour, R. Brydson, K. J. Roberts, M. Ghadiri, H. Stitt, Effect of mill type on the size reduction and phase transformation of gamma alumina. *Chem. Eng. Sci.* **134**, 774–783 (2015). [doi:10.1016/j.ces.2015.06.004](https://doi.org/10.1016/j.ces.2015.06.004)
11. A. K. Garg, Fine alpha alumina particles, their production and use. European patent EP0554908A1 (1993).
12. H. Ma, A. Krell, Synthesis and processing of nano- $\alpha\text{-Al}_2\text{O}_3$ powders. *Key Eng. Mater.* **206–213**, 43–46 (2001).
13. B. Michen, C. Geers, D. Vanhecke, C. Endes, B. Rothen-Rutishauser, S. Balog, A. Petri-Fink, Avoiding drying-artifacts in transmission electron microscopy: Characterizing the size and colloidal state of nanoparticles. *Sci. Rep.* **5**, 9793 (2015). [doi:10.1038/srep09793](https://doi.org/10.1038/srep09793) [Medline](#)
14. B. Michen, S. Balog, B. Rothen-Rutishauser, A. Petri-Fink, D. Vanhecke, TEM sample preparation of nanoparticles in suspensions: Understanding the formation of drying artefacts. *Imaging Microsc.* **6**, 39 (2014).

ACKNOWLEDGMENTS

We thank A. Schlüter from the electron microscopy department at the Max-Planck-Institut für Kohlenforschung for recoding the TEM images. **Funding:** Supported by the Max-Planck-Institut für Kohlenforschung.

13 February 2020; accepted 6 May 2020
Published online 29 May 2020
[10.1126/science.abb0948](https://doi.org/10.1126/science.abb0948)

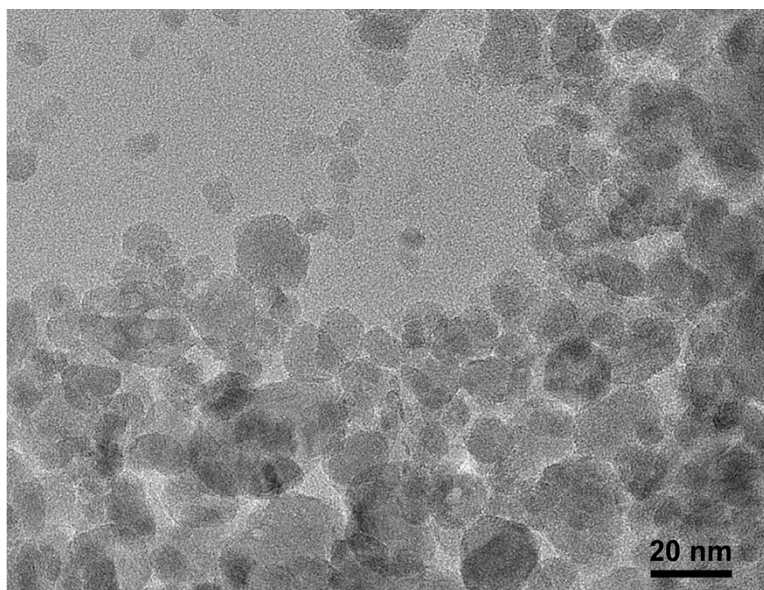


Fig. 1. TEM image of corundum NPs derived by ball milling of boehmite. TEM analysis was performed on a Hitachi HF-2000 microscope with a cold field-emission cathode at a maximum acceleration voltage of 200 kV. Samples were prepared by suspending the specimen in ethanol, followed by ultrasonication for 20 min, placing a drop of sample suspended in ethanol on the TEM grid, and allowing it to dry under ambient conditions.



Students present their talents as they pitch their innovations at the third annual HBCU Making & Innovation Showcase in Washington, D.C., in February 2020.

Diverse students display inventions at HBCU showcase

Innovators from historically black institutions address challenges being faced by their communities

By **Andrea Korte**

A social media platform for online education, a smart cane that warns visually impaired users of objects along their path, and a mobile app that connects customers to restaurants eager to distribute surplus food were just a few of the inventions that student innovators from historically black colleges and universities displayed at a recent showcase.

The American Association for the Advancement of Science, with support from the National Science Foundation, hosted the third annual HBCU Making & Innovation Showcase in February 2020 in Washington, D.C. The event, held in conjunction with AAAS's and NSF's Emerging Researchers National Conference, brought together 80 students and faculty members from HBCUs for 2 days of interactive workshops and training on invention and entrepreneurship. The conference culminated with 18 teams of undergraduate and graduate students from 13 colleges and universities pitching their innovations.

"Entrepreneurial thinking and an invention mindset are critical to address the challenges that affect communities around the world, and in our own backyards," said Neela White, a project director in AAAS's Diversity, Equity, and Inclusion program. "As we look at the shifting demographics in the country, we must be intentional about ensuring that innovation ecosystems are fully inclusive."

The showcase was inspired by a similar effort held during the White House's 2015 National Week of Making, which celebrated the maker movement: individuals and groups harnessing technology and creativity to turn their innovative ideas into reality. Quincy Brown, a AAAS STEM program director at the time, proposed the HBCU Making & Innovation Showcase to the National Science Foundation in 2017 and hosted the first showcase in 2018.

"The showcase is designed to provide an opportunity for students to display the talent and innovation already at HBCUs, show how these students are addressing challenges specific to their communities, heighten the awareness of the individuals and resources that support an inclusive innovation ecosystem, and further develop students' skills," White said.

Each of the teams created an innovative solution to a problem in its community related to one of 17 United Nations Sustainable Development Goals, which include quality education, clean water and sanitation, affordable and clean energy, and "no poverty." The teams designed and built prototypes to address a pressing challenge in their community. Each participant then filmed a video of their prototype to demonstrate and pitch their solutions to expert judges, who selected the top three projects at the showcase.

The winning team—Nicolette Barrieffe, a student at Clark Atlanta University, and Stephen Seymour and Leoul Tilahun, both from Morehouse College—tackled the U.N. goal of climate action with a network of communication devices intended for use in disasters and other emergency situations. The Guardian Network allows individuals to report their status when internet connectivity and cellular services are not available using long-range communication networks. The network also features a dashboard to allow emergency operators and first responders to keep track of those in need and allocate assistance.

The February presentation by Barrieffe, Seymour, and Tilahun was not the first time they took part in the showcase. The trio competed in the showcase's inaugural competition in 2018 as first-year college students, placing second.

"Based on their success, they wanted to do bigger and better projects," said Ayodeji Oyesanya, who served as the team's faculty adviser in 2018 and 2020. Oyesanya runs the Morehouse College

AAAS NEWS & NOTES

MakerSpace Exploration Center and advised five teams participating in this year's showcase.

In the break between showcases, Seymour and Barrieffe co-founded MakeWay, a student organization dedicated to designing innovative projects that now counts more than 60 members. "This year, I got the gang back together," said Seymour. "But now of course, you see the world in a different way. Our perspectives are more matured and developed."

"The trio returned to the showcase determined to win," noted Oyesanya, and they achieved their goal.

As an adviser, Oyesanya seeks to encourage students to solve problems on their own. "I find it remarkable how these students can make such wonderful projects and such creative projects and innovative projects with what could be perceived as limited resources," he said. "Because we are an HBCU, we don't have the resources and the million-dollar budget of a lot of predominantly white institutions. But these students have learned not to allow that to stagnate them and to still be able to compete with any of these schools in the country."

Added White, "I was blown away by the talent of all of the students, by the level of knowledge they brought and how much more they were willing to absorb. It was just really great."

White was also impressed with how strongly students connected with the sustainability themes to inform their projects. "They are driven by social good; they are driven by wanting to do something good for their region, for the global well-being," she said.

"Our number one goal whenever we work on a project is that it's culturally relevant," said Barrieffe. When Seymour's family was affected by Hurricane Dorian, which hit the Bahamas in 2019, the team realized that emergency communications would be an important place to focus their efforts, she said.

In addition to the competition, students and faculty attending the showcase took part in workshops on subjects such as technology transfer, the business of entrepreneurship, and career pathways in the innovation sector.

Another session focused on collaboration and teamwork in innovation.

"It's really hard to make new things happen when you're by yourself," said Diana Yousef, who, along with her business partner Huda Elasaad, spoke during a session titled "The Power of Teams: From Invention to Entrepreneurship."

Yousef and Elasaad are the chief executive officer and chief technology officer, respectively, of change:WATER Labs, which has developed a new type of toilet for use in places lacking sanitation infrastructure. Their "iThrone" uses a waste-evaporating material and a urine-powered bio-battery to dehydrate human waste as a way to replace flushing in places with no sewage plumbing.

The pair were also the first ever to be selected together as co-ambassadors for the AAAS-Lemelson Invention Ambassadors Program, which seeks to increase understanding of the critical role of invention in improving quality of life and to encourage a new and diverse generation of inventors.

During the showcase, Yousef and Elasaad emphasized to participants that working with a team allows everyone to contribute their strength. Yousef trained as a lab scientist and holds an MBA, while Elasaad is a water system engineer. As the pair traversed the innovation-to-market pathway, they brought varied abilities to the effort. "We cover different parts of that journey," said Yousef. "We have these different skill sets that all come together, and it makes it all possible."

Yousef later reflected on her experiences as a woman and daughter of immigrants in the innovation sphere.



Nicolette Barrieffe and Stephen Seymour explain their winning innovations.

"There's definitely a lot of discounting that happens," she said. "To see what AAAS and this initiative are doing to promote serious innovators and scientists who don't look like the scientists that everybody imagines is really amazing."

Knowing the importance of seeing oneself represented in a community, White made sure the event featured speakers from a diverse range of backgrounds, including minorities, women, and HBCU graduates who are now leaders in tech and entrepreneurship.

"It was very relatable for all of the students," White said.

Participants also had the opportunity to be inspired by each other.

"I learned a lot from my peers, especially those who have been seasoned with years of experience," said Barrieffe. "It's definitely powerful, especially as a woman of color in STEM."

Under the leadership of AAAS STEM Program Director Iris R. Wagstaff, the current principal investigator of the NSF grant that supports the showcase, new strategic partnerships and collaborations have been formed, expanding the showcase and providing additional resources and training to students. Social media and online platforms also extend the showcase experience.

The conversation continues online with the newly launched Emerging Researchers National Conference in STEM LinkedIn community for participants of the ERN conference and HBCU Making & Innovation Showcase, White said. The online platform means that tech and innovation leaders can reach a much broader cross section of students, including those whose ideas might not yet be fully fledged to take part in a showcase, she noted. Yet, the community has already become a hub for resource-sharing, regular online workshops, and ongoing conversation for students unable to convene in person.

Innovation, after all, is an ongoing, iterative process, White said, adding that students who attended the 2020 showcase were "very eager to go back to their campuses and continue working on these innovations."

The students from the winning team also plan to continue pursuing innovation and entrepreneurship.

Seymour, for instance, is concerned that the Bahamas—his home country—and the broader Caribbean region lack STEM education resources. "There's a huge demand for it," said Seymour, adding that he hopes to get involved with "finding ways to push the envelope on how we utilize our resources" to encourage students to explore STEM disciplines from an early age. Focusing on the social capacity of innovation is another factor driving his interest.

"Sometimes the greatest ideas are just the ones that have the impact," Seymour said. "If you can't really help or change someone's way of living, how impactful is that?"

RESEARCH

IN SCIENCE JOURNALS

Edited by Michael Funk

EVOLUTIONARY BIOLOGY

Primate kin recognition

Kin selection, the selection of behaviors that benefit kin, is key to social and cooperative behavior. Primates can recognize kin through heritable facial traits. The question is whether this behavior is incidental from shared genes or if it is instead subject to selection. Charpentier *et al.* studied the behavior of mandrills, nonhuman primates that live in enormous groups composed of matriline, in which daughters stay with the mother

and sons disperse. Using sophisticated artificial intelligence, they graded facial resemblances and correlated them with the social interactions of maternal and paternal half-siblings. The authors found evidence of selection for kin recognition and discuss differences between maternal and paternal half-siblings and the intricacies resulting from different social settings and the selective forces involved. —ABR *Sci. Adv.* 10.1126/sciadv.aba3274 (2020).

Recognition of kin influences the social interactions of mandrills (shown), social primates that form large groups.

SEPARATIONS

Zeolites that prefer alkynes

Alkenes such as ethylene and propene must be separated from alkynes before they can be converted in polymers. Drawbacks in current methods, such as hydrogenation of alkynes producing unwanted alkanes, has spurred interest in sorption separation methods. Zeolites have generally been inefficient, given the similar sizes and volatilities of the molecules. Chai *et al.* incorporated atomically dispersed divalent transition metal cations into faujasite zeolite and found that the nickel-containing analog efficiently removed alkynes from

olefins through chemoselective binding at open nickel(II) sites. At ambient conditions in the presence of water and carbon dioxide, the zeolites retained separation selectivities of 100 and 92, respectively, for acetylene over ethylene and propyne over propylene for 10 adsorption-desorption cycles. —PDS

Science, this issue p. 1002

CORONAVIRUS

Alternative hosts and model animals

The severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2) pandemic may have originated in bats, but how

it made its way into humans is unknown. Because of its zoonotic origins, SARS-CoV-2 is unlikely to exclusively infect humans, so it would be valuable to have an animal model for drug and vaccine development. Shi *et al.* tested ferrets, as well as livestock and companion animals of humans, for their susceptibility to SARS-CoV-2 (see the Perspective by Lakdawala and Menachery). The authors found that SARS-CoV-2 infects the upper respiratory tracts of ferrets but is poorly transmissible between individuals. In cats, the virus replicated in the nose and throat and caused inflammatory pathology deeper in the respiratory tract, and airborne transmission did occur between

pairs of cats. Dogs appeared not to support viral replication well and had low susceptibility to the virus, and pigs, chickens, and ducks were not susceptible to SARS-CoV-2. —CA

Science, this issue p. 1016;
see also p. 942

BLOOD-BRAIN BARRIER

Transport vehicle for CNS therapeutics

Delivering therapeutics to the brain is complicated by the presence of the blood-brain barrier. Kariolis *et al.* and Ullman *et al.* developed a transport vehicle (TV) for central nervous system (CNS) delivery of therapeutics

and tested it in mice and monkeys. The TV was obtained by binding an antibody fragment of the human immunoglobulin G1 to the transferrin receptor expressed in brain endothelial cells. Fusing the TV to the anti- β -secretase antibody resulted in high expression of the antibody in the CNS in mice and monkeys. In a mouse model of lysosomal storage disorder, peripheral delivery of iduronate 2-sulfatase fused to the TV had therapeutic effects. The TV might be effective for delivering therapeutics in neurological disorders. —MM

Sci. Transl. Med. **12**, eaay1359, eaay1163 (2020).

OCEAN CIRCULATION

Changing forces in midstream

The intensity and frequency of the strongest cyclones east of Taiwan have increased over the past several decades as the climate has warmed. Zhang *et al.* found that one result of this trend has been the strengthening of Kuroshio current transport off the coast of Japan. The Kuroshio, like its Atlantic counterpart the Gulf Stream, is a surface current that moves huge volumes of warm water from low latitudes to high ones. As strong Pacific cyclones have become stronger, they have increased the amount of energy contained in cyclonic mesoscale ocean eddies and decreased that of anticyclonic ones. This in turn has increased the transfer of energy to the Kuroshio as eddies move into the current, providing a feedback between



Tropical cyclones, such as Typhoon Maria in 2018, seen in a satellite image, appear to be strengthening the Kuroshio current in the North Pacific.

climate warming and ocean heat transport. —HJS

Science, this issue p. 988

ORGANIC CHEMISTRY

Stitching alkynes into bryostatin 3

The bryostatin family of marine natural products has been explored for a wide variety of pharmaceutical applications but remains challenging to source. The general structure comprises a macrocycle that contains three smaller, six-membered rings. Bryostatin 3 is distinguished by the added complexity of a fourth, fused lactone ring. Trost *et al.* report a convergent synthesis of this complex molecule, taking advantage of alkyne coupling reactions to stitch together three main fragments and asymmetric dihydroxylation and propargylation reactions to set stereochemistry. —JSY

Science, this issue p. 1007

CANCER

Profiling tumor bacteria

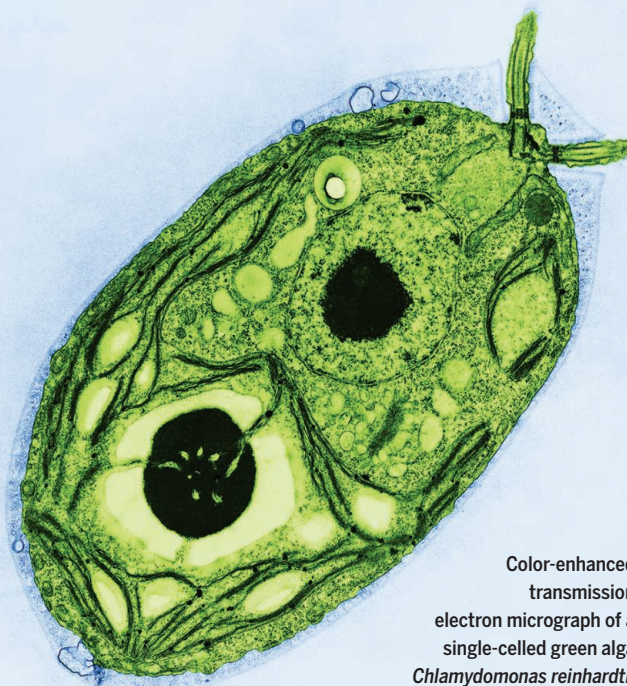
Bacteria are well-known residents in human tumors, but whether their presence is advantageous to the tumors or to the bacteria themselves has been unclear. As an initial step toward addressing this question, Nejman *et al.* produced an exhaustive catalog of the bacteria present in more than 1500 human tumors representing seven different tumor types (see the Perspective by Atreya and Turnbaugh). They found that the bacteria within tumors were localized within both cancer cells and immune cells and that the bacterial composition varied according to tumor type. Certain biologically plausible associations were identified. For example, breast cancer subtypes characterized by increased oxidative stress were enriched in bacteria that produce mycothiol, which can detoxify reactive oxygen species. —PAK

Science, this issue p. 973; see also p. 938

Science, this issue p. 973; see also p. 938

IN OTHER JOURNALS

Edited by **Caroline Ash**
and **Jesse Smith**



Color-enhanced transmission electron micrograph of a single-celled green alga *Chlamydomonas reinhardtii*

BIOLOGICAL MEMBRANES

Dissecting uneven complex distribution

Eukaryotic cells contain heterogeneous membranes that vary in curvature, lipid and protein composition, and cellular context. Filling in molecular details for such cellular structures is a great challenge for structural biologists. Wietrzynski *et al.* used in situ cryo-electron tomography to capture views of the native thylakoid membranes of the single-celled green alga *Chlamydomonas reinhardtii*. Advances in data collection and processing permitted identification of specific membrane-associated complexes such as photosystems I and II, adenosine triphosphate synthase, and thylakoid-associated ribosomes. Two-dimensional projections of the membrane surface revealed a sharp compositional transition between appressed membranes (those that directly face another membrane) and nonappressed membranes. This technique should enable study of how these membranes are organized at both the cellular and molecular levels and how they react to different light conditions. —MAF

eLife **9**, e53740 (2020).

MEDICAL GENETICS

Genetic variant takes the pressure off

The identification of rare genetic variants that protect carriers from a specific disease can provide a launch point for studies of disease biology and therapy. In a search for genes that affect

the risk of developing the common eye disease glaucoma, Tanigawa *et al.* examined data from more than 500,000 individuals represented in UK and Finnish biobanks. They found that missense and nonsense variants in *ANGPTL7*, the gene encoding angiopoietin-related protein 7, which is a member of

a protein family implicated in angiogenesis, were associated with lower intraocular pressure and reduced risk of glaucoma, including a variant that reduced risk by 34%. Consistent with a role in glaucoma, ANGPTL7 is expressed in the trabecular meshwork, a tissue that drains fluid (aqueous humor) from the eye. —PAK

PLOS Genet. **16**, e1008682 (2020).

IMMUNOLOGY

Commensals produce a gut reaction

Commensal bacteria in the small intestine are known to help shape immune responses. Less well understood is whether the microbiota residing in the stomach are similarly immunoregulatory. Satoh-Takayama *et al.* report that commensal bacteria can indeed regulate group 2 innate lymphoid cell (ILC2) homeostasis in the stomach. Furthermore, *Helicobacter pylori*, a pathogen responsible for gastritis and gastric cancer in humans, rapidly induces stomach ILC2 proliferation and activation in mice in an interleukin-7 (IL-7)– and IL-33–dependent manner. IL-5 production by ILC2 triggers B cell secretion of immunoglobulin A, which plays a role in *H. pylori* containment. ILC2 and commensal bacteria in the stomach may therefore serve as targets for various gastrointestinal disorders. —STS

Immunity **52**, 635 (2020).

NEUROSCIENCE

Dopamine circuits facilitate fear learning

To ensure survival, powerful mechanisms of pain sensation and fear learning have evolved in animals. One important fear-learning center is the amygdala of the mammalian brain. Dopamine neurons in the midbrain ventral tegmental area (VTA) project into several brain regions, including the amygdala. Using in vivo optogenetics and optrode recordings, as well as

anatomic circuit tracing and cFos imaging, Tang *et al.* studied the role of this dopamine pathway in the formation of fear memory. The authors found that a few VTA dopamine neurons projecting into the mouse amygdala were activated when a mild shock was administered to the animal's foot. Dopamine release thus facilitates the formation of a fear memory. —PRS

J. Neurosci. **40**, 3969 (2020).

ORGANIC CHEMISTRY

Epoxidizing arenes

Although the metabolism of aromatic rings often involves epoxide formation, this reaction has proven challenging to translate into applications in synthetic chemistry. Siddiqi *et*

al. now report a straightforward sequential protocol for arene epoxidation: A photochemical cycloaddition with a triazoline reagent first dearomatizes the arene and then a manganese catalyst introduces the oxygen in the epoxide ring. Oxidative removal of the triazoline completes the process. Minor modification of the conditions for benzenes also cleanly produced ring-expanded oxepines from naphthalene and quinoline reactants. —JSY

J. Am. Chem. Soc. **10.1021/jacs.0c02724** (2020).

PHYSICS

Tuning the Chern number

The material MnBi_2Te_4 is an antiferromagnet that can, in

thin-film form, support interesting topological states. Ge *et al.* investigated the transport properties of MnBi_2Te_4 films 7 to 10 layers thick in the presence of a magnetic field. The authors showed that the films were in the so-called Chern insulator state, exhibiting quantized conductance that could not be explained by the ordinary quantum Hall effect. The thinner, seven-layer samples had a Chern number of 1, whereas the 9- and 10-layer samples were in an even more exotic state with a Chern number of 2. The tuning of the Chern number with sample thickness was also supported by numerical calculations. —JS

Natl. Sci. Rev. **10.1093/nsr/nwaa089** (2020).

GRASSLAND

Move the fences

The sensitive alpine meadow and steppe systems of the Tibetan Plateau have experienced serious degradation over the past half-century. To restore these habitats, an extensive system of wire fences has been erected across the region; some have been in place for 30 years. Fences can protect plants from immediate grazing by livestock, but they limit connectivity for other organisms, interrupt trophic dynamics, and artificially divide landscapes. Sun *et al.* used a large-scale meta-analysis to determine whether these fences have been effective for restoration, how they affect wildlife, and what effect they have on human populations on the basis of interviews with local herdsman. Fences that had been in place for short to medium periods of time were able to increase aboveground vegetative biomass for both meadows and steppe. However, long-term fencing decreased plant growth and diversity, with negative ecosystem impacts. In addition, fences inhibited the movement of three focal mammal species—Tibetan gazelles, yaks, and donkeys—which increased their grazing impact on unfenced regions. The herders perceived fences as not only preventing their ability to use traditional grazing practices but also as being ineffective overall. Fences can be useful tools but only when they are transitional and impermanent. —SNV

Sci. Bull. **10.1016/j.scib.2020.04.035** (2020).



The high-altitude grasslands of Tibet have become badly degraded, but fencing out key grazers, such as these Tibetan antelopes, is not a long-term solution.

ALSO IN *SCIENCE* JOURNALS

Edited by Michael Funk

MAGNETISM**A structurally ordered spin glass**

Spin glasses that form in disordered materials such as magnetic alloys have locally varying magnetic patterns, and their spin relaxation occurs over time scales spanning many orders of magnitude. Kamber *et al.* used spin-polarized scanning tunneling microscopy to image the magnetism on the (0001) surface of thick, single-crystal films of neodymium as a function of temperature and magnetic field. Despite the lack of structural disorder, they found a spectral distribution of degenerate magnetic wave vectors, or *Q* states, that exhibited spatio-temporal variation. In this spin-*Q* glass, pockets of nearly degenerate spin spiral states formed with varying periodicity. Ab initio electronic structure coupled to atomistic spin dynamics calculations suggests that the double hexagonal closed packed crystal structure in neodymium drove strongly frustrated magnetism that created these pockets. —PDS

Science, this issue p. 966**FOREST ECOLOGY****Shifting forest dynamics**

Forest dynamics are the processes of recruitment, growth, death, and turnover of the constituent tree species of the forest community. These processes are driven by disturbances both natural and anthropogenic. McDowell *et al.* review recent progress in understanding the drivers of forest dynamics and how these are interacting and changing in the context of global climate change. The authors show that shifts in forest dynamics are already occurring, and the emerging pattern is that global forests are tending toward younger stands with faster turnover as old-growth forest with stable dynamics are dwindling. —AMS

Science, this issue p. 964**SYNTHETIC BIOLOGY****Electronic control of designer cells**

There is increasing interest in using designer cells to produce or deliver therapeutics. Achieving direct communication between such cells and electronic devices would allow precise control of therapies. Krawczyk *et al.* describe a bioelectronic interface that uses wireless-powered electrical stimulation of cells to promote the release of insulin (see the Perspective by Brier and Dordick). The authors engineered human β cells to respond to membrane depolarization by rapidly releasing insulin from intracellular storage vesicles. A bioelectronic device that incorporates the cells can be wirelessly triggered by an external field generator. When subcutaneously implanted in type 1 diabetic mice, the device could be triggered to restore normal blood glucose levels. —VV

Science, this issue p. 993;
see also p. 936**CORONAVIRUS****Safe vaccine development**

The coronavirus disease 2019 (COVID-19) pandemic has prompted accelerated vaccine development in the hope that predicted mortality rates can be reduced and population, or herd, immunity achieved. These measures could result in eventual eradication of the disease. In a Perspective, Graham discusses the need to ensure that vaccines are safe and do not aggravate coronavirus infection. Based on lessons learned from past vaccines, various steps need to be taken to ensure that expedited vaccine development is accompanied in parallel by safety assessments to prioritize the most effective candidates. —GKA

Science, this issue p. 945**IMMUNOLOGY****Granzyme A lights a fire**

Cytotoxic T cells and natural killer cells use several strategies to kill infected or transformed cells. One such pathway entails the delivery of a family of serine proteases called granzymes to target cells through perforin-mediated pores to induce a form of programmed cell death called apoptosis. Zhou *et al.* show that granzyme A cleaves and activates gasdermin B (GSDMB), a central player in the highly inflammatory cell death process known as pyroptosis (see the Perspective by Nicolai and Raulet). GSDMB expression was highly expressed in some tissues and could be up-regulated by interferon- γ . Enforced expression of GSDMB in cancer cells enhanced tumor clearance in a mouse model, suggesting that this pathway may be a target for future cancer immunotherapies. —STS

Science, this issue p. 965;
see also p. 943**INVASIVE SPECIES****Exotic plants reduce carbon sequestration**

Invasive exotic plants have become a major problem worldwide, with transformational effects on the composition and function of ecosystems. In a multifactorial experiment in New Zealand, Waller *et al.* show that exotic plants accelerate carbon loss from soils through their interactions with invertebrate herbivores and soil biota (see the Perspective by Urcelay and Austin). They built 160 mini-ecosystems in the field, manipulating interactions among plants, invertebrate herbivores, and soil biota. Key biological and abiotic responses were measured to quantify the relative contribution and interactions of the components of each community, revealing the potential of invasive plants to

influence and suppress carbon sequestration through biotic interactions. —AMS

Science, this issue p. 967;
see also p. 934**PEPTIDE SYNTHESIS****Fully synthetic whole proteins in reach**

Solid-phase peptide synthesis of homogeneous peptides longer than about 50 amino acids has been a long-standing challenge because of inefficient coupling and side reactions. Hartrampf *et al.* used an automated chemistry platform to optimize fast-flow peptide synthesis and were able to produce fully synthetic single-domain proteins (see the Perspective by Proulx). The targets included proinsulin and enzymes such as barnase and a version of HIV-1 protease containing multiple noncanonical amino acids. Refolded peptides were nearly indistinguishable from recombinant proteins, and the synthesized enzymes had activity close to that of their ribosomally synthesized counterparts. This method will enable fast, on-demand synthesis of small proteins with a vastly expanded pool of precursor amino acids. —MAF

Science, this issue p. 980;
see also p. 941**CORONAVIRUS****Coronavirus in nonhuman primates**

We urgently need vaccines and drug treatments for coronavirus disease 2019 (COVID-19). Even under these extreme circumstances, we must have animal models for rigorous testing of new strategies. Rockx *et al.* have undertaken a comparative study of three human coronaviruses in cynomolgus macaques: severe acute respiratory syndrome–coronavirus (SARS-CoV) (2002), Middle East respiratory syndrome (MERS)–CoV

(2012), and SARS-CoV-2 (2019), which causes COVID-19 (see the Perspective by Lakdawala and Menachery). The most recent coronavirus has a distinct tropism for the nasal mucosa but is also found in the intestinal tract. Although none of the older macaques showed the severe symptoms that humans do, the lung pathology observed was similar. Like humans, the animals shed virus for prolonged periods from their upper respiratory tracts, and like influenza but unlike the 2002 SARS-CoV, this shedding peaked early in infection. It is this cryptic virus shedding that makes case detection difficult and can jeopardize the effectiveness of isolation. —CA

Science, this issue p. 1012;
see also p. 942

ICE SHEETS

A rapid retreat

Are the rates at which we observe ice shelves shrinking today representative of how fast they shrank in the past? Dowdeswell *et al.* report observations of the Antarctic seafloor that reveal the presence of submarine grounding-zone wedges on the Larsen continental shelf (see the Perspective by Jakobsson). The authors interpret these ridges as being caused by the tidal rise and fall of the ice shelf at the grounding line, which squeezes the underlying sediments when it rests on the seafloor. From this, they calculated that ice shelf retreat at this location about 14,000 years ago was at times as much as 100 times as fast as the average over the past 10,000 years. —HJS

Science, this issue p. 1020;
see also p. 939

CELL BIOLOGY

Promoting cancer from the Golgi

Signaling through insulin-like growth factor 1 receptor (IGF-1R) promotes cancer progression, but therapies targeting IGF-1R signaling have performed poorly. Rieger *et al.* found that

phosphorylation-dependent trafficking of IGF-1R promotes aggressive cancer cell behaviors (see the Focus by Crudden and Girnita). Ligand stimulation induced the translocation of IGF-1R to the Golgi, where it enhanced cell migration before being recycled to the plasma membrane. These findings may explain why targeting plasma membrane-localized IGF-1R has not been clinically effective. —AMV

Sci. Signal. **13**, eaba3176, eabb7887
(2020).

MACROPHAGES

Targeting overzealous macrophages

Intestinal homeostasis relies on the maintenance of a complex set of interactions between intestinal microbiota and the intestinal immune system. Pathogens that colonize the gut disrupt these interactions and promote intestinal inflammation. Corbin *et al.* used the mouse pathobiont *Helicobacter hepaticus*, which causes inflammation akin to human inflammatory bowel disease, to study the role of intestinal macrophages in driving inflammation. Using this model, the authors found that the transcription factor interferon regulatory factor 5 (IRF5) was a critical regulator of macrophage inflammatory potential. Deletion of IRF5 rendered mice resistant to *H. hepaticus*-driven intestinal inflammation. IRF5 and molecules upstream of IRF5 may be potential drug targets in the treatment of human inflammatory bowel disease. —AB

Sci. Immunol. **5**, eaax6085 (2020).

REVIEW SUMMARY

FOREST ECOLOGY

Pervasive shifts in forest dynamics in a changing world

Nate G. McDowell*, Craig D. Allen, Kristina Anderson-Teixeira, Brian H. Aukema, Ben Bond-Lamberty, Louise Chini, James S. Clark, Michael Dietze, Charlotte Grossiord, Adam Hanbury-Brown, George C. Hurtt, Robert B. Jackson, Daniel J. Johnson, Lara Kueppers, Jeremy W. Lichstein, Kiona Ogle, Benjamin Poulter, Thomas A. M. Pugh, Rupert Seidl, Monica G. Turner, Maria Uriarte, Anthony P. Walker, Chonggang Xu

BACKGROUND: Forest dynamics arise from the interplay of chronic drivers and transient disturbances with the demographic processes of recruitment, growth, and mortality. The resulting trajectories of vegetation development drive the biomass and species composition of terrestrial ecosystems. Forest dynamics are changing because of anthropogenic-driven exacerbation of chronic drivers, such as rising temperature and CO₂, and increasing transient disturbances, including wildfire, drought, windthrow, biotic attack, and land-use change. There are widespread observations of increasing tree mortality due to changing climate and land use, as well as observations of growth stimulation of younger forests due to CO₂ fertilization. These antagonistic processes are co-occurring globally, leaving the fate of future forests uncertain. We examine the implications of changing forest demography and its drivers for both future forest management and forecasting impacts of global climate forcing.

ADVANCES: We reviewed the literature of forest demographic responses to chronic drivers and transient disturbances to generate hy-

potheses on future trajectories of these factors and their subsequent impacts on vegetation dynamics, with a focus on forested ecosystems. We complemented this review with analyses of global land-use change and disturbance datasets to independently evaluate the implications of changing drivers and disturbances on global-scale tree demographics. Ongoing changes in environmental drivers and disturbance regimes are consistently increasing mortality and forcing forests toward shorter-statured and younger stands, reducing potential carbon storage. Acclimation, adaptation, and migration may partially mitigate these effects. These increased forest impacts are due to natural disturbances (e.g., wildfire, drought, windthrow, insect or pathogen outbreaks) and land-use change, both of which are predicted to increase in magnitude in the future. Atmospherically derived estimates of the terrestrial carbon sink and remote sensing data indicate that tree growth and potentially recruitment may have increased globally in the 20th century, but the growth of this carbon sink has slowed. Variability in growth stimulation due to CO₂ fertilization is evident globally, with observations

and experiments suggesting that forests benefit from CO₂ primarily in early stages of secondary succession. Furthermore, increased tree growth typically requires sufficient water and nutrients to take advantage of rising CO₂. Collectively, the evidence reveals that it is highly likely that tree mortality rates will continue to increase, whereas recruitment and growth will respond to changing drivers in a spatially and temporally variable manner. The net impact will be a reduction in forest canopy cover and biomass.

OUTLOOK: Pervasive shifts in forest vegetation dynamics are already occurring and are likely to accelerate under future global changes, with consequences for biodiversity and climate forcing. This conclusion is robust with respect to the abundant literature evidence and our global assessment of historical demographic changes, but it also forms the basis for hypotheses

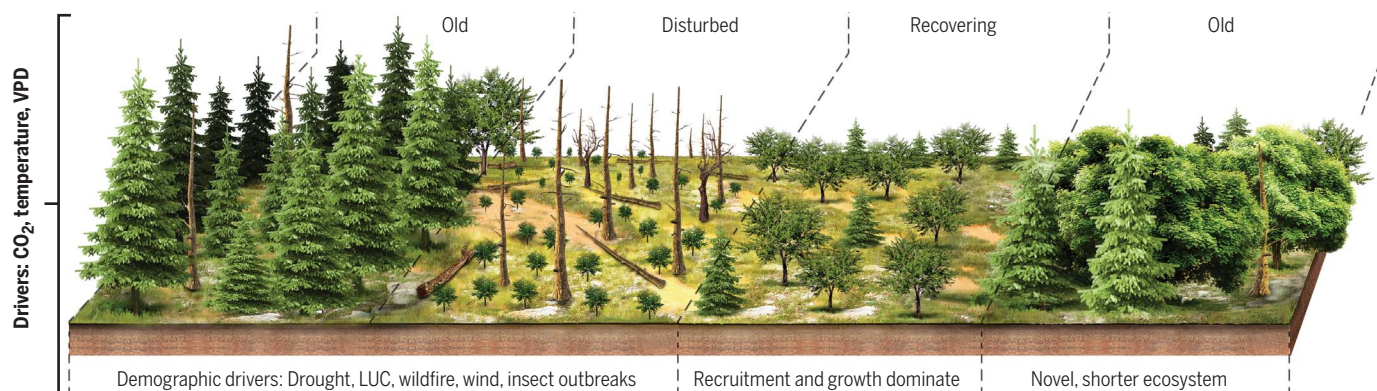
ON OUR WEBSITE

Read the full article at <https://dx.doi.org/10.1126/science.aaz9463>

regarding the patterns and processes underlying the shifts in forest dynamics. These hypotheses will be directly testable using emerging terrestrial and satellite-based observation networks. The existing evidence and newly made observations provide a critical test of Earth system models that continue to improve in their ability to simulate forest dynamics and resulting climate forcing. Ultimately, forest managers and natural resource policies must confront the consequences of changing climate and disturbance regimes to ensure sustainable forests and accrue their associated benefits. ■

The list of author affiliations is available in the full article online.

*Corresponding author. Email: nate.mcdowell@pnsl.gov
Cite this article as N. G. McDowell *et al.*, *Science* **368**, eaaz9463 (2020). DOI: [10.1126/science.aaz9463](https://doi.org/10.1126/science.aaz9463)



A conceptual diagram of the components of forest dynamics and the disturbances that drive them. In the far-left panel, a mature ecosystem is responsive primarily to localized mortality, and the primary drivers of demography are chronically changing variables such as CO₂, temperature, and vapor pressure deficit (VPD). In the next panel, the system is disturbed by fire, insect outbreak, or another large-scale perturbation that removes most of the overstory trees,

and species adapted to rapid postdisturbance recruitment become established. In the third panel, recruitment and growth dominate demographic processes, with mortality increasing over time as competition leads to self-thinning. In the last panel, a mature ecosystem is dominated by species that have replaced the original community in response to chronic environmental changes, leading to a novel ecosystem.

REVIEW

FOREST ECOLOGY

Pervasive shifts in forest dynamics in a changing world

Nate G. McDowell^{1*}, Craig D. Allen², Kristina Anderson-Teixeira^{3,4}, Brian H. Aukema⁵, Ben Bond-Lamberty⁶, Louise Chini⁷, James S. Clark⁸, Michael Dietze⁹, Charlotte Grossiord¹⁰, Adam Hanbury-Brown¹¹, George C. Hurtt⁷, Robert B. Jackson¹², Daniel J. Johnson¹³, Lara Kueppers^{11,14}, Jeremy W. Lichstein¹⁵, Kiona Ogle¹⁶, Benjamin Poulter¹⁷, Thomas A. M. Pugh^{18,19}, Rupert Seidl^{20,21}, Monica G. Turner²², Maria Uriarte²³, Anthony P. Walker²⁴, Chonggang Xu²⁵

Forest dynamics arise from the interplay of environmental drivers and disturbances with the demographic processes of recruitment, growth, and mortality, subsequently driving biomass and species composition. However, forest disturbances and subsequent recovery are shifting with global changes in climate and land use, altering these dynamics. Changes in environmental drivers, land use, and disturbance regimes are forcing forests toward younger, shorter stands. Rising carbon dioxide, acclimation, adaptation, and migration can influence these impacts. Recent developments in Earth system models support increasingly realistic simulations of vegetation dynamics. In parallel, emerging remote sensing datasets promise qualitatively new and more abundant data on the underlying processes and consequences for vegetation structure. When combined, these advances hold promise for improving the scientific understanding of changes in vegetation demographics and disturbances.

The interplay of vegetation demography—recruitment, growth, and mortality—with environmental conditions and disturbances drives forest dynamics of biomass, function, and species composition (see Box 1 for definitions). In old-growth forests that approximate steady-state demographics, the recruitment, growth, and mortality of trees are approximately balanced; in contrast, rapid recruitment often follows widespread disturbance-induced mortality (1). Vegetation dynamics may now be changing because the environmental context in which plant demography and disturbances interact is shifting with anthropogenic change. The interaction between episodic forest disturbances, such as windthrow or wildfire, and chronically changing drivers, such as rising temperature, vapor pressure deficit (VPD), and CO₂, together with land-use change (LUC) (2), leads to both compounding and antagonistic impacts that alter demographic rates (3), with consequences for terrestrial biogeochemical cycles and climate (4, 5). Understanding the drivers of vegetation

dynamics is thus critical for accurate prediction of global terrestrial biogeochemistry under future conditions (6).

The impacts of global change on forest demographic rates may already be materializing. In mature ecosystems, tree mortality rates have doubled throughout much of the Americas and in Europe over the past four decades (7–9). Simultaneously, global carbon budgets indicate either a growing or constant terrestrial carbon sink (10–12), which implies increased or constant vegetation production rates (13, 14). However, satellite evidence suggests that forests might be switching from a CO₂ fertilization-dominated period to a VPD-dominated period (15). Terrestrial greening indices indicate a shift from a CO₂-driven increase in greenness in the late 20th century to a VPD-driven decrease in the past decade (16). Thus, increasing mortality due to anthropogenic changes and potentially increasing or stable growth and recruitment due to CO₂ fertilization (5) represent opposing processes that are co-occurring globally, leaving the fate of future forests uncertain.

In addition to changing vegetation dynamics in intact or relatively undisturbed forests, episodic disturbances are tending to be larger, more severe, and in some regions more frequent under global climate change (17–20). Similarly, the rates and types of LUC vary widely (21) but have, on average, increased globally in the past few centuries (2, 22, 23). Thus, at the global scale, disturbances and LUC have likely amplified tree mortality beyond what is suggested by the doubling of background mortality rates in undisturbed forests (7–9). Current understanding of the net balance of tree losses (mortality) and gains (recruitment and growth) under a changing environment characterized by more-extreme drivers and disturbances is limited, preventing prediction of whether recruitment and growth can balance increased mortality rates in the future.

To evaluate whether environmental changes and increasing disturbances are causing globally widespread shifts in vegetation demography, we reviewed global observations of recruitment, growth, and mortality of forests and woodlands. Our expert-derived compilation of the state-of-the-art knowledge on vegetation dynamics, their drivers, and disturbances, allowed us to address four questions: (i) Is there evidence for shifts in demography over recent decades? (ii) What physiological and disturbance-mediated processes underlie these demographic shifts? (iii) What are the potential consequences of disturbance-mediated changes in demography for climate forcing? (iv) How can global predictions of future vegetation dynamics best be improved?

Evidence for changing drivers and disturbances and their impact on demography

Determining the impacts of changing drivers on demography is difficult given the lack of global observation platforms. However, evidence abounds from individual published studies on the drivers and their impacts on plant communities, and new modeling and observational efforts now enable a more complete picture of disturbances and forest demography (24–26). In this section, we first examine whether there are global trends in stand ages and test the sensitivity of the stand-age distribution to changes in disturbance rate

¹Pacific Northwest National Laboratory, Richland, WA 99354, USA. ²U.S. Geological Survey, Fort Collins Science Center, New Mexico Landscapes Field Station, Los Alamos, NM 87544, USA.

³Conservation Ecology Center, Smithsonian Conservation Biology Institute, Front Royal, VA 22630, USA. ⁴Forest Global Earth Observatory, Smithsonian Tropical Research Institute, Republic of Panama. ⁵Department of Entomology, University of Minnesota, St. Paul, MN 55108, USA. ⁶Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD 20740, USA. ⁷Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA. ⁸Nicholas School of the Environment, Duke University, Durham, NC 27708, USA. ⁹Department of Earth and Environment, Boston University, Boston, MA 02215, USA. ¹⁰Swiss Federal Institute for Forest, Snow and Landscape Research WSL, 8903 Birmensdorf, Switzerland. ¹¹Energy and Resources Group, University of California, Berkeley, Berkeley, CA 94720, USA. ¹²Department of Earth System Science, Woods Institute for the Environment, and Precourt Institute for Energy, Stanford University, Stanford, CA 94305, USA. ¹³School of Forest Resources and Conservation, University of Florida, Gainesville, FL 32611, USA. ¹⁴Division of Climate and Ecosystem Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. ¹⁵Department of Biology, University of Florida, Gainesville, FL 32611, USA. ¹⁶School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ 86001, USA. ¹⁷Biospheric Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. ¹⁸School of Geography, Earth and Environmental Sciences, University of Birmingham, B15 2TT Birmingham, UK. ¹⁹Birmingham Institute of Forest Research, University of Birmingham, B15 2TT Birmingham, UK.

²⁰Department of Forest and Soil Sciences, University of Natural Resources and Life Sciences Vienna, 1180 Vienna, Austria. ²¹School of Life Sciences, Technical University of Munich, 85354 Freising, Germany. ²²Department of Integrative Biology, University of Wisconsin–Madison, Madison, WI 53706, USA. ²³Department of Ecology, Evolution and Environmental Biology, Columbia University, New York, NY 10027, USA. ²⁴Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. ²⁵Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.

*Corresponding author. Email: nate.mcdowell@pnnl.gov

Box 1. Vegetation dynamics definitions.

We focus on three main plant demographic processes: recruitment, growth, and mortality. Recruitment (including reproduction) determines the seedling and sapling composition of a plant community after disturbance (152). Growth from sapling to mature plant results in development of mature forests and includes competitive processes. Mortality is a key rate controlling carbon storage and species composition in a plant community and is a dominant demographic rate during a pulse disturbance (153, 154).

Abiotic drivers. Physical factors that cause changes in demography and that respond to global change or to disturbances, such as light, CO₂, soil moisture, humidity, temperature, etc.

Biotic drivers. Biological factors that may drive changes in demography, such as pathogens, insects, herbivores, or competition with other individuals.

Chronic environmental change. Persistently changing drivers of demographic rates. These drivers have a nonstable and directional trajectory, such as rising CO₂, temperature, and VPD.

Demographic rate. Any individual-, population-, or community-level parameter that affects the age and/or size structure of a population or community, including rates of recruitment, growth, and death.

Demographic driver. An abiotic or biotic factor that, when undergoing a change itself, also leads to change(s) in one or more demographic rates.

Disturbance. The destruction of live plant biomass in a discrete event (155, 156).

Disturbance regimes. Spatial and temporal characteristics of disturbances in a landscape over a long time period, including frequency, return interval, duration, intensity, severity, and size.

Growth. The rate of biomass production over time, at the individual or ecosystem scale (i.e., net primary production in grams of carbon per square meter per year).

Land-use and land-cover change. Anthropogenic shifts in forms of cultivation or in vegetation cover due to, for example, forestry or conversion of woodlands to crop ecosystems.

Mortality. Defined herein as the complete loss of a plant's ability to reproduce and ultimately the loss of cellular metabolism.

Recruitment. The rates of transition of plants from one size class to another (typically in units of individuals per square meter per year). Recruitment results from the birth and growth of individuals. Herein, we consider recruitment from the stage of seed dispersal through seedling growth into the sapling stage.

Self-thinning. Reduction in the number of live plants within a stand, occurring via competition for resources.

Vegetation dynamics. The net outcome of the interplay between disturbances and vegetation demographic rates.

using global datasets on LUC (27) and non-LUC (25, 28) disturbances. We subsequently draw upon the wealth of published studies on changes in forest demographics and their drivers to investigate the potential changes leading to global stand-age trends. Ultimately, the combination of our global estimates and the large literature base allows us to generate testable hypotheses regarding trends and impacts of the drivers of forest demographics.

Is disturbance changing forest demography at the global scale?

We reanalyzed the Land-use Harmonization (LUHv2) dataset (27) with respect to forest age, revealing that the area of young forest stands (here defined as stands younger than 140 years old) resulting directly from LUC (conversion of forest to nonforest) or wood harvest (forest retention with reduction of biomass and age) has increased from 4.8 million km² in 1900 to 12.5 million km² in 2015 (or from 11.3 to 33.6% of forest area) (Fig. 1A). The results were insensitive to assumptions regarding the link of disturbance likelihood to stand age (Fig. 1A). These forest stand-age distributions exhibit different trajectories in different regions. Tropical forests have pro-

gressively lost old-growth area owing to LUC over the course of the 20th century (Fig. 2A, black dashed line). Wood harvest has increased from a minor driver of tropical forest age distribution in 1900 to a major one in 2015 (difference between solid and dashed lines). The split between deforestation and shifting cultivation drivers is broadly consistent with a satellite-based analysis for the period 2001–2015 (29). Temperate and Mediterranean forest ages are strongly influenced by wood harvest, which has made old-growth forests increasingly scarce in these regions. LUC has had minimal influence on stand age in boreal forests, but wood harvest has substantially shifted boreal forest age distribution toward young growth.

In reality, old-growth forests are made scarcer by more than just LUC and wood harvest (Figs. 1A and 2), they are threatened by other disturbances that have shifted landscapes from old to young growth-dominated stands (14), such as wildfire (29), windthrow (30), and biotic outbreaks (31). To address these additional disturbances, we integrated recent observation-based estimates of non-LUC disturbance for closed-canopy forests (25, 28) with LUC from LUHv2 to obtain a first-principles

estimate of the combined effect of human and natural disturbances on forest age structure (Fig. 1B). A twofold increase in non-LUC disturbance rates over the period 2015–2050 would result in a substantial increase in the fraction of young forests (Fig. 1, B and C). Thus, realistic shifts in disturbance rates can substantially affect the age structure of forests in the future. As discussed below, such an increase in disturbance rate is consistent with the magnitude of changes observed or predicted in individual ecosystems.

Notably, calculations based on the Global Forest Age Dataset (GFAD) v1.1 (14, 32) yielded 16.5 million km² of old-growth forest and 26.3 million km² of young forest (32), which differs from what is shown in Fig. 1, B and C. This disparity is likely attributable to consideration of different forest types (closed-canopy forests versus all forests) and to differences in definitions of stand size and age used in inventories versus those used in satellite-based estimates.

Chronically changing drivers Atmospheric CO₂

Atmospheric CO₂ has risen more than 125 parts per million (ppm) since the industrial revolution (11) and is projected to rise an additional 50 to 200 ppm by 2100. Higher CO₂ increases leaf-level water-use efficiency, and rising CO₂ has positive but uncertain feedbacks on plant demographic rates (Fig. 3, A and B). Maturation and seed production can be accelerated under elevated CO₂ (33); however, seedling growth is not always stimulated by CO₂ (34). Recruitment response to rising CO₂ is variable (35, 36). Forest inventory and tree-ring studies show limited evidence for CO₂ fertilization of growth (37–43), potentially because of the overwhelming influence of increasing drought and nutrient limitations (44). Ecosystem-scale CO₂ enrichment experiments in young forests suggest a 30% gain in decadal biomass increment (45), but experiments in mature forests have found minimal growth stimulation (46, 47). This is consistent with evidence for an initially strong CO₂-related growth stimulation in young forests that decreases with tree age and size (39) perhaps due to nutrient (7, 48) and hydraulic path-length limitations (49).

A limited number of studies suggest that elevated CO₂ causes increased mortality or no change in mortality. Mortality rates of saplings during experimental drought were not mitigated by elevated CO₂ (50, 51), while accelerated self-thinning due to CO₂ fertilization-induced stand density increases may lead to higher mortality (6, 52, 53) (Fig. 3B). The latter process would be consistent with increases in recruitment at large scales. Because tree mortality is dominated by large size classes [i.e., (54)] (for details see section on size-related mortality below), faster growth via CO₂ fertilization may

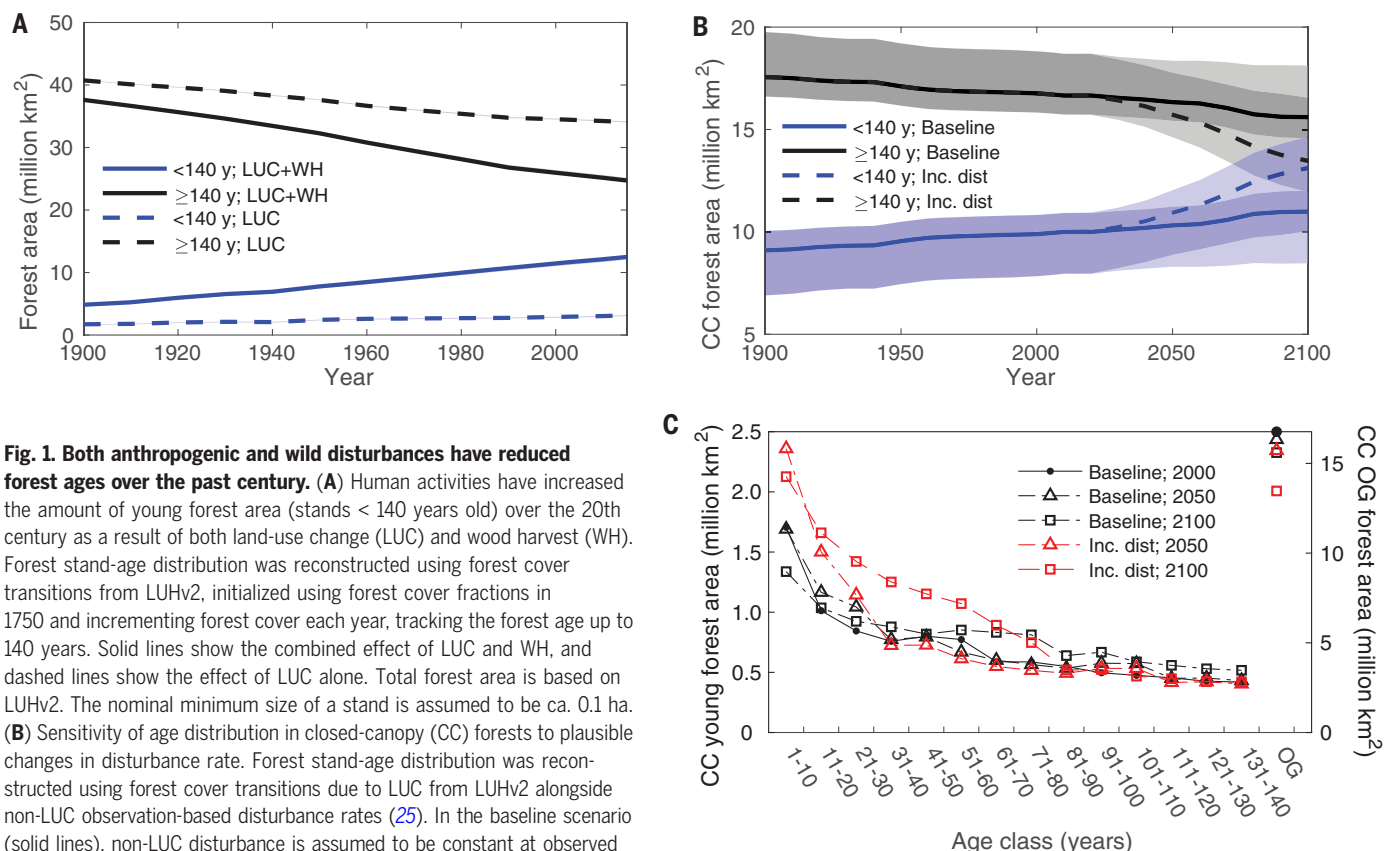


Fig. 1. Both anthropogenic and wild disturbances have reduced forest ages over the past century. (A) Human activities have increased the amount of young forest area (stands < 140 years old) over the 20th century as a result of both land-use change (LUC) and wood harvest (WH). Forest stand-age distribution was reconstructed using forest cover transitions from LUHv2, initialized using forest cover fractions in 1750 and incrementing forest cover each year, tracking the forest age up to 140 years. Solid lines show the combined effect of LUC and WH, and dashed lines show the effect of LUC alone. Total forest area is based on LUHv2. The nominal minimum size of a stand is assumed to be ca. 0.1 ha. **(B)** Sensitivity of age distribution in closed-canopy (CC) forests to plausible changes in disturbance rate. Forest stand-age distribution was reconstructed using forest cover transitions due to LUC from LUHv2 alongside non-LUC observation-based disturbance rates (25). In the baseline scenario (solid lines), non-LUC disturbance is assumed to be constant at observed 2001–2014 values throughout. In the incremented distribution scenario ("Inc. dist"; dashed lines), disturbance rates are incremented linearly to 200% of the 2001–2014 values over the period 2015–2050 and held constant at that level thereafter. The underlying LUC scenario is Global Change Assessment Model Representative Concentration Pathway 3.4 (GCAM RCP 3.4), which includes land-based mitigation for CO₂ emissions. Results are presented for CC forests only (25), which is why total forest area is lower in (B) than in (A), as non-LUC disturbance rate information is not currently available for open-canopy forests. The shaded areas in (A) and (B) indicate the effect of assuming that disturbances are five times more likely to affect young forests than old-growth

forests, or vice versa, as opposed to an even probability across ages (solid lines). The apparent large dampening of this assumption in (A) versus (B) is primarily due to the different y-axis scales. **(C)** Changes in the disturbance regime propagate through forest age structure at decadal time scales. CC young (<140 years old) forest area is shown on the left-hand y axis. Old-growth (OG; >140 years old) forest area is shown on the right-hand y axis (same units) and refers to the data points in the upper right-hand corner of the panel. Scripts used and additional methods can be accessed at https://github.com/pugtam/AgeClassReconst_rel.git.

expose trees to size-related mortality risks earlier (7). Such CO₂-induced increases in mortality may be global (55). Furthermore, faster growth is often associated with lower wood density (56), rendering fast growing trees more susceptible to high winds. Thus, future CO₂ fertilization could increase recruitment, growth, and mortality (Fig. 4B), although there is considerable uncertainty about these effects.

Temperature and vapor pressure deficit

Temperature and VPD are rising globally and will continue to rise into the future (57). Both temperature and VPD can have impacts on demographic rates. Rising temperature forces an exponential rise in VPD, which prompts stomatal closure and limits photosynthesis, leading to lower growth, higher mortality (58), and reduced regeneration (59) and ultimately driving community shifts (60, 61). These ob-

servations are consistent with hydraulic theory, which suggests that as VPD rises, potential maximum tree height declines (62) (Fig. 4). This results from the dependency of water transport limitations on tree size (49) that are exacerbated by elevated VPD (Fig. 4), making short stature advantageous with rising VPD. Because most plants cannot reduce their size (beyond limited reductions in leaf area or crown dieback), forests respond through increased mortality of large plants, which are replaced by smaller ones (62), as has been observed in many studies (26, 54). While rising air temperature may also increase respiratory carbon loss, leaving less carbon for growth (63), warming in wetter and cooler regions may actually stimulate reproductive output, recruitment, and growth (3, 64, 65). Changes in temperature and VPD also can produce asynchrony in floral and pollinator phenology (66) and can reduce cold stratification (67),

both of which reduce seed abundance (68), and negatively affect recruitment (69, 70). Sapling mortality is accelerated by elevated temperature (70, 71), but recruitment has increased in moist areas (72). Thus, rising temperature and VPD may be beneficial in cooler or wetter areas, but most evidence suggests negative impacts on plant demographic rates (Figs. 3, C and D, and 4).

Changing disturbance regimes Droughts

Droughts are anticipated to increase in frequency, duration, and severity globally (Fig. 3, E and F) and are more stressful to plants owing to increases in temperature, VPD, and associated water loss (57). Drought can directly cause tree death or indirectly lead to mortality through associated increases in insect or pathogen attack (51). Hydraulic failure and carbon starvation remain the most likely, mutually

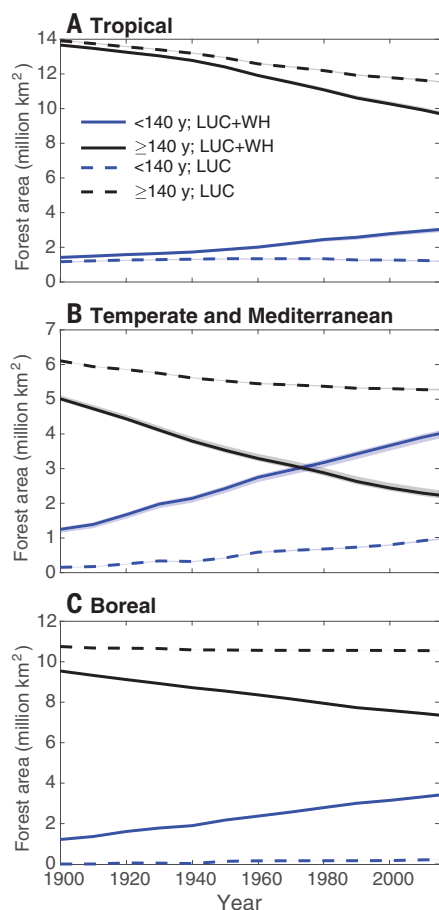


Fig. 2. Human activities have increased the amount of young forest area irrespective of biome. As in Fig. 1A, but broken down by biome (157): (A) tropical, (B) temperate and Mediterranean, and (C) boreal.

inclusive, underlying physiological mechanisms for drought-induced mortality (73), and both processes are likely to increase tree susceptibility to biotic agents (74). Evidence suggests that drought-induced mortality occurs more rapidly under warmer conditions (51, 71). Consistent with these empirical results, models suggest far greater mortality of temperate conifer trees in the future (75). Reproductive output is often reduced by drought [but see (64)], which, combined with drought impacts on seedling survival, leads to reduced recruitment (76). However, growth was relatively stable across a drought in Amazonia (77) while mortality increased. Thus, like rising temperature and VPD, it appears that drought may increase mortality regardless of location, while having variable impacts on recruitment and growth (Fig. 3F).

Land-use change

LUC and forest management have reduced vegetation stature and biomass and have altered species composition, with profound consequences for forest dynamics (Figs. 1A

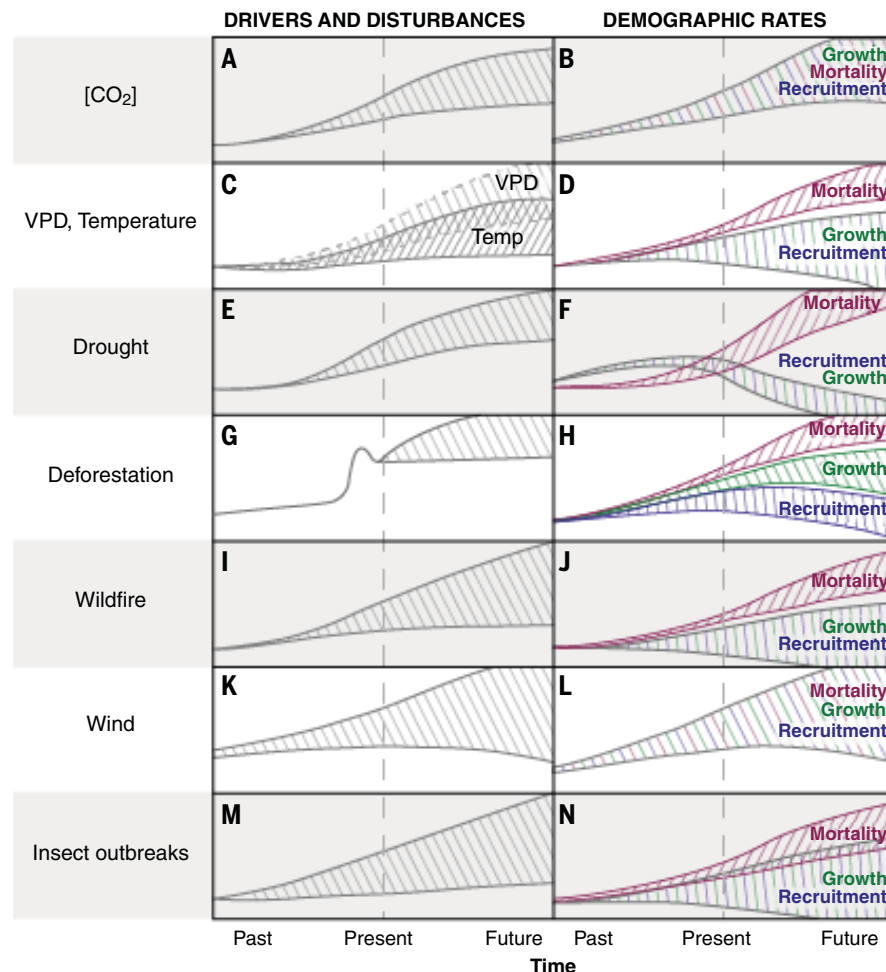


Fig. 3. Drivers, disturbances, and demographics are changing both historically and into the future.

A graphical summary of the literature evidence of changing drivers and disturbances (left-hand column) and subsequent demographic rates (right-hand column). Shown are the chronically changing drivers (A and B) CO₂ and (C and D) VPD and temperature, as well as the more transient disturbances of (E and F) drought (low precipitation), (G and H) deforestation, (I and J) wildfire, (K and L) wind, and (M and N) insect outbreaks. Each driver or disturbance's corresponding demographic responses (shown as carbon fluxes per unit area over time) are shown in the corresponding right-hand panels.

and 3, G and H). Today's global vegetation biomass stocks may amount to only ~50% of their potential because of LUC (78). Wood harvest and shifting cultivation are the land-use activities primarily responsible for the conversion from primary to secondary vegetation cover and associated demographic shifts (2). In systems that return to wild vegetation or to managed forest after human clearing, demographic rates are typically accelerated. The increased resource availability after forest removal facilitates establishment of early successional species, reduces species diversity (79, 80), and triggers a transition to younger, smaller plants (81). Post-deforestation recruitment is often prolific even in the absence of management (82). Globally, the recovery of harvested forests and abandoned agricultural land, along with establishment of new planta-

tions, has resulted in younger forests (Fig. 1A), with associated reductions in tree size and biomass (83). Such post-deforestation recruitment may be limited by elevated VPD or drought, as is the case with recruitment after all-natural disturbances. Overall, the net effect of historical LUC and wood harvest has resulted in a substantial loss of forest area, along with altered demographic rates, leading to younger, shorter, less diverse ecosystems (Fig. 3H).

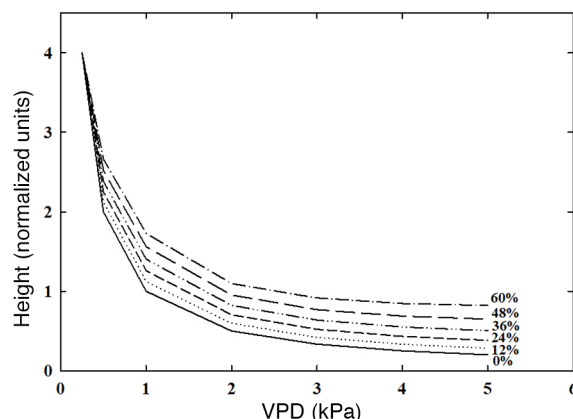
Wildfire

Wildfire is increasing in many forests worldwide (84) (Fig. 3I), although human management of landscapes has led to wildfire suppression in some biomes (85). Given sufficient fuel, burned area increases exponentially with aridity (86), and future fire frequencies may exceed those documented over the past

Fig. 4. Rising VPD forces declines in potential plant stature.

Predictions of plant height in response to rising VPD from the hydraulic corollary to Darcy's law. The equation is $h = (A_s \times k_s \times \Delta\Psi) / (G \times A_l \times \text{VPD})$, where h is height, A_s is sapwood area, k_s is specific conductivity, $\Delta\Psi$ is the leaf-to-soil water potential gradient, G is stomatal conductance, and A_l is leaf area (53). The different lines represent different levels of acclimation of A_s , k_s , $\Delta\Psi$, G , and A_l , all allowed to adjust simultaneously from 0 to 60% of their initial values.

In the case of G , it is assumed to decrease because of rising atmospheric CO_2 . Acclimation can help, but not completely mitigate, the impact of rising VPD on plant stature.



10,000 years (87). Increased fire activity causes increased mortality and potentially higher recruitment and growth of either preexisting or newly introduced species, but rates of recruitment and growth may be slowed under climate warming. Forests characterized by stand-replacing fire regimes are dominated by obligate seeders and typically have effective seedling recruitment (88). However, high-severity and high-frequency fires can reduce recruitment by reducing seed supply through the repeated and severe loss of reproductively mature vegetation (89), and high-frequency fires can cause recruitment losses via direct mortality of the seedbank, seedlings, and saplings (90), which is worsened by elevated VPD (59). Woody species that can resprout after fire, including shrubs that suppress tree regeneration (59), may be favored by increased fire frequency and severity. Increased fire severity results in high tree mortality in forests historically adapted to low-severity fires, and subsequent recruitment and growth may be slow or absent, resulting in conversion of forests to low-biomass ecosystems (91). Thus, wildfire can result in higher demographic rates, although rising temperature and VPD can negatively affect recruitment and growth (Fig. 3J).

Windthrow

Windthrow from cyclonic storms represents the dominant natural disturbance in coastal forests across the globe (92). Cyclonic storms are expected to increase in frequency, wind velocities, and precipitation intensity (93) (Fig. 3K), resulting in more extreme flooding that promotes tree instability. Windthrow also results from convective thunderstorms and topographically mediated winds, and warming is expected to increase the frequency of atmospheric conditions conducive to severe thunderstorms (94). Canopy damage and

whole-tree mortality are the most immediate impacts of windthrow (95) (Fig. 3L). Storm-induced mortality is greatest for larger trees (96), and the loss of large canopy trees during wind disturbance favors growth of surviving trees (96, 97) and advances regeneration, recruitment of early successional species (98), or resprouting of trees broken by wind (99). Depending on the resprouting or seeding capacity of surviving species, wind damage may either slow or accelerate succession (100). We note that storms may also be associated with lightning, which may be a prominent cause of large-tree mortality (101). Thus, windstorms should result in changes in all three demographic rates, although with large uncertainty at the global scale (Fig. 3L).

Biotic agents

Biotic disturbances from insects, insect-pathogen complexes, and other biotic agents have been increasing in frequency, severity, and extent in recent decades (17, 31, 102) (Fig. 3M). Such trends reflect a changing climate (103), altered land use (104), and introductions of nonindigenous insects and pathogens (105). Climate change is expected to further amplify biotic disturbances (106), in part through enhanced host vulnerability (Fig. 3M). However, shifts in frequency or dampening of disturbance regimes could also emerge (107), leading to some uncertainty in outbreak dynamics under future conditions (Fig. 3M). Whereas insects and associated pathogens are globally widespread, lianas, or vines that use other plants as host structures, are increasing in abundance and are thought to be causing increasing mortality in the tropics (7, 108).

Response of insects and pathogens to climate change is likely to increase plant mortality (4), with variable impacts on growth and recruitment (Fig. 3N). Tree mortality can result from girdling of the phloem and xylem by

bark beetles (74) and from repeated defoliation events that exhaust the capacity of trees to recover (109). Tree mortality during outbreaks is usually partial at the stand level because many biotic agents preferentially attack trees of specific size or health classes or are host-specific (16). Suppressed, smaller trees and nonhost tree species may survive and grow rapidly when released from competition for resources (110, 111). Thus, similar to many other disturbances, mortality increases while recruitment and growth show variable responses to biotic disturbances, including a dependency on post-disturbance temperature, VPD, and drought.

On size and age demographics

The combination of LUC, disturbances, and chronic drivers is likely to have already shifted forests to younger and shorter stands, with these impacts increasing under expected future changes in drivers and disturbances (Fig. 1, A to C). These results are consistent with our review of the literature (Fig. 3). Large trees are the most susceptible to die from LUC-induced forest fragmentation (112, 113), drought (26), rising temperature or VPD (54, 62) (Fig. 4), windthrow (114, 115), biotic attacks (116), and lightning (101), with variable size impacts of fire (117). The abundance of size-dependent mortality drivers and disturbances should logically push stands toward younger and smaller distributions of trees and shorter-statured species assemblages (118).

There are exceptions to the pattern of climate drivers and disturbances reducing tree height and stand age. Non-stand-replacing fires that kill smaller trees but spare the larger, older trees will shift forests toward larger size distributions. Similarly, on occasions when droughts preferentially kill younger but fast-growing trees, subsequent size distribution and rate of carbon accumulation would be affected. Rising CO_2 and increased precipitation in some areas also counter the general decrease in size, because they may lead to faster growth and hence taller trees (119). Thus, the antagonistic drivers promoting larger trees (e.g., rising CO_2) and smaller trees (e.g., rising VPD, increasing disturbances) co-occur, but the general pattern of decreasing size and younger ages reveals that processes driving down size and age (Figs. 1 to 4) are dominant globally.

Mitigation of demographic-disturbance impacts

The literature patterns suggest that most drivers and disturbances will increase tree mortality now and in the future, with variable effects on recruitment and growth (Fig. 3). This supposition becomes uncertain, however, when we consider multiple feedbacks that can mitigate the changes in forest demography induced by chronically changing drivers and disturbance regimes. These processes include

acclimation, adaptation, migration, and compensatory mechanisms of resource use. With global change, forests will be influenced by a combination of phenotypic plasticity [i.e., acclimation (120)], adaptation to novel biotic and abiotic stresses (121), and the ability to migrate as conditions change (122). Failure to acclimate, adapt, or migrate—including failure due to human infrastructure (123)—could lead to recruitment and growth reductions and local extinctions. Plants have demonstrated acclimation of phenology, seed longevity, and metabolic processes to single and/or multiple stressors (124–127). Acclimation and adaptation will likely depend on an array of factors including genetic variation, fecundity, dispersal, population size, and environmental variability (120). Many tree species have migrated in response to past climatic cycles but at rates slower than the current pace of climate change (128). Regarding resource use, reductions in stand density as a result of increased mortality or reduced recruitment should allow greater resource availability to surviving individuals and therefore subsequently higher growth and survival rates (129). Such stand resource mechanisms can manifest at the landscape scale, as most disturbances are patchy (130), and the size, shape, and arrangement of surviving forest patches can play a key role in recovery of the disturbed landscape (20). Taken together, the mitigating factors can play a substantial role in buffering the impacts of changing drivers on plant survival, but it remains unclear whether these factors will enable trees to keep pace with ongoing climate change (50, 120). Ultimately, the uncertainty surrounding future demographic rates shown in Fig. 3 is partially due to the influence of these mitigating factors.

Consequences for community assembly and for climate forcing

The widespread shift in vegetation dynamics begets questions regarding consequences for community assembly and climate forcing. Hydraulic theory suggests that under rising VPD, functional traits of high conductance, low stature, and low leaf area should best enable survival, all of which are characteristics of pioneer, shrub, and weed species (62). Consistent with this theory, diversity (e.g., species richness) temporarily increases post-disturbance for many systems, as short-statured, opportunistic species invade (131). If forest communities shift toward trait assemblages better suited to the new disturbance regime, such shifts may confer some resistance to future disturbances (131, 132). Alternatively, if disturbance regimes shift faster than recruitment, growth, and subsequent community assembly can respond, resistance to future disturbances will likely decline.

Climate forcing responds to changing vegetation dynamics in complex ways. Changes

in forest disturbance regimes and vegetation dynamics can affect climate via biogeochemical, hydrological, and land-surface energy budgets (133). Reductions in biomass result in a loss of carbon to the atmosphere despite younger, shorter stands often having higher gross photosynthesis; this is due to the loss of carbon through decomposition of necromass, which is a particularly large flux from mortality of older, larger trees, such as those in old-growth forests (134), and reduced landscape-mean carbon storage under an intensified disturbance regime (135). The time required for an ecosystem to reacheive the same live carbon storage after disturbance can be decades to centuries, particularly if the disturbance cycle is increased, thus the net effect of the biomass loss is increased CO₂ to the atmosphere and hence greater climate forcing. This impact may be mitigated by increased carbon uptake due to CO₂ fertilization (119) or enhanced recruitment. Calculations of the terrestrial carbon sink from atmospheric inversions indicate that the sink grew over recent decades (12) in part because of increased leaf area (13), which is consistent with increased recruitment and growth. However, evidence suggests that forests are switching from a CO₂ fertilization-dominated period to a VPD-dominated period (15, 16), despite sustained high gross photosynthesis at the global scale (136). The increased mortality throughout much of the terrestrial biosphere (7–9) further minimizes potential carbon storage through enhanced biomass loss. Ultimately, the terrestrial contribution to climate forcing through carbon uptake and release results from the antagonistic process of rising CO₂ and forest recovery from LUC, which enhance the carbon sink, and rising VPD and disturbances that reduce the carbon sink.

Changing vegetation dynamics also influence regional and global surface energy budgets and hydrological cycles. Disturbances frequently shift albedo of ecosystems from darker to lighter, resulting in a decline in radiative forcing through less light absorption (137). The rate of recruitment after disturbance influences the temporal period of this negative feedback (138). The impact of changing vegetation dynamics on the water cycle is particularly complex. Evaporation from canopies shifts as stands become taller, because taller trees transpire less (per unit leaf area) than smaller trees (49), but larger trees often have better rooting access to water sources and have greater total leaf area. The net effect of disturbance is a transient decrease in evaporative loading to the atmosphere along with albedo shifts, causing a feedback of decreasing precipitation downwind (139, 140). Ultimately, carbon storage is at least transiently reduced by disturbances,

with mixed impacts on the water and energy budgets.

The path to improved prediction

Changes in global drivers (temperature, CO₂, and VPD) and disturbances (including LUC, drought, wildfire, windstorms, and insect outbreaks) should all force forests toward shorter, younger, lower-biomass ecosystems. This trend is supported by hydraulic theory (62) (Fig. 4) and by abundant empirical evidence demonstrating a consistent increase in mortality across the global spectrum of drivers and disturbances and variable, but often declining, recruitment and growth (Fig. 3). While the bulk of the evidence points to reduced plant stature owing to changing drivers, large uncertainty remains in the magnitude and slope of demographic trajectories in the future (Fig. 3). Given these trajectories, and the large uncertainties around them, what are the critical next steps to allow improved global prediction? Continued long-term observations (both on the ground and remotely sensed) are essential to reveal the patterns of demographic responses to drivers and disturbances. Likewise, manipulative experiments are needed that alter conditions such as CO₂ or drought to provide cause-and-effect understanding of the interactions among mechanisms of demographic responses. However, for global-scale prediction of responses and climate consequences, we need to mainstream insights from observations and experiments into Earth system models (ESMs).

ESMs simulate the exchange of fluxes between the atmosphere, land, and ocean and stores of carbon, water, and energy; the land-surface modules of ESMs simulate vegetation. ESMs have made great progress in simulating land use, disturbances, and demography, including representation of wildfire (141), drought-induced mortality (142), and cohort-age structured models that enable representation of succession and associated shifts in physiological traits (6). The global Coupled Model Intercomparison Project CMIP6 now includes a dedicated model intercomparison activity focused on the effects of changes of land use on carbon and climate (143). Advances in remote sensing and forest inventory integration are enhancing global datasets of forest structure (144) and age (32) that can be used in model initialization, data assimilation benchmarking, and sensitivity analyses (Fig. 1, A to C). These advancements set the stage for developments in ESMs, such as the prediction of disturbances and demographic rate responses under climate and LUC scenarios.

The newest generation of ESMs uses size or age-structured approaches to explicitly model demography in the Earth system (6), which should ultimately enable model-based representation of observed shifts in age structure

(e.g., Fig. 1). However, representation of vegetation demographic rates remains relatively simplistic. Simulation of growth responses to global change requires model refinement in light capture, belowground water and nutrient acquisition, and responses of respiration to temperature (6). Recruitment, including reproduction and dispersal, is the most undeveloped demographic process in ESM simulations. Reproductive allocation is invariant with plant functional type (PFT), and seed is assumed to mix evenly throughout a grid cell [but see (145)]. Environmental constraints to PFT establishment are derived from prior distributions of major taxa, and while recruitment rates can be influenced by light or space availability, they are not responsive to temperature, CO₂, or soil moisture (146, 147). Simplistic dispersal assumptions are typically either overly permissive or overly restrictive. Improvements in representing recruitment under global change are critical for improving predictions of vegetation dynamics. These advancements will require data synthesis and additional data collection to support PFT-specific, environmentally sensitive parameterizations of regeneration processes, such as reproductive allocation; effective dispersal; seedling establishment, survival, and growth; and post-disturbance recovery strategies (e.g., serotiny and resprouting).

Disturbance-induced mortality is better developed for landscape-scale models than for ESMs. ESM modeling of disturbance-induced mortality exists for wildfire and drought (141, 142), although considerable challenges remain to reliably represent both disturbances globally, while ESMs are underdeveloped for wind and insect mortality. To our knowledge, only one ESM currently represents canopy damage (148); this causes ESMs to potentially underestimate the impacts of drought and wind, as both disturbances cause lagged tree mortality associated with canopy loss years after the inciting event (149, 150). As for insects, there have been prescriptive studies examining the impact of insect outbreaks on land processes within ESMs, but no ESM has yet explicitly considered the interaction between plant defense and insect population dynamics for prediction of large-scale insect-induced tree mortality. For predicting wildfire, models should be sensitive to both fuels and climate interactions and represent spatial patterns of burn severity, because the burn mosaic strongly influences postfire vegetation dynamics (141). Next-generation demographic models are evolving to include explicit, mechanistic representations of drought-associated mortality, including carbon starvation and hydraulic failure (151). The evaluation of new hydraulics models (151) for prediction of mortality is an essential next step. Ultimately, model formulations that include environmentally sen-

sitive, PFT-specific processes compatible with the cohort-based approach are likely to provide the best compromise between process detail and parsimony and are therefore most likely to capture changes in large-scale forest dynamics under future conditions.

Outlook

Forest vegetation dynamics are already strongly influenced by global change (Fig. 1) and will continue to be affected in the future (Figs. 1 to 4) by changes in land use, chronic drivers such as CO₂ and VPD, and increasing frequency and severity of transient disturbances such as windthrow, wildfire, and insect outbreaks. Effects on forests are driven largely by consistent increases in tree mortality from these drivers, and variable responses of recruitment and growth depending on stand age, disturbance type, and geographic location (Fig. 3). The consequences of changing demographics suggest an increasing constraint in terrestrial carbon storage due, at least, to the consistent increase in mortality. Any declines in recruitment or growth, especially when disturbance-recovery cycles are disrupted, will exacerbate this carbon-cycle constraint. Shifts in other terrestrial radiative forcing terms such as energy and water budgets are also likely. Although well supported by the literature, data, and sensitivity analysis (Fig. 1), the trends in Fig. 3 represent hypotheses to be tested by the next generation of observational platforms, both terrestrial and spaceborne. Forest management must ultimately confront the elevated mortality and uncertainty in recruitment and growth when considering options for sustaining the societal benefits of forests into the future.

REFERENCES AND NOTES

1. M. G. Turner, T. G. Whithy, D. B. Tinker, W. H. Romme, Twenty-four years after the Yellowstone Fires: Are postfire lodgepole pine stands converging in structure and function? *Ecology* **97**, 1260–1273 (2016). doi: [10.1890/15-1585.1](https://doi.org/10.1890/15-1585.1); pmid: [27349102](https://pubmed.ncbi.nlm.nih.gov/27349102/)
2. G. C. Hurtt *et al.*, Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Clim. Change* **109**, 117–161 (2011). doi: [10.1007/s10584-011-0153-2](https://doi.org/10.1007/s10584-011-0153-2)
3. K. J. Anderson-Teixeira *et al.*, Altered dynamics of forest recovery under a changing climate. *Glob. Change Biol.* **19**, 2001–2021 (2013). doi: [10.1111/gcb.12194](https://doi.org/10.1111/gcb.12194); pmid: [23529980](https://pubmed.ncbi.nlm.nih.gov/23529980/)
4. M. Reichstein *et al.*, Climate extremes and the carbon cycle. *Nature* **500**, 287–295 (2013). doi: [10.1038/nature12350](https://doi.org/10.1038/nature12350); pmid: [23955228](https://pubmed.ncbi.nlm.nih.gov/23955228/)
5. R. Seidl *et al.*, Forest disturbances under climate change. *Nat. Clim. Change* **7**, 395–402 (2017). doi: [10.1038/nclimate3303](https://doi.org/10.1038/nclimate3303); pmid: [28861124](https://pubmed.ncbi.nlm.nih.gov/28861124/)
6. R. A. Fisher *et al.*, Vegetation demographics in Earth System Models: A review of progress and priorities. *Glob. Change Biol.* **24**, 35–54 (2018). doi: [10.1111/gcb.13910](https://doi.org/10.1111/gcb.13910); pmid: [28921829](https://pubmed.ncbi.nlm.nih.gov/28921829/)
7. N. McDowell *et al.*, Drivers and mechanisms of tree mortality in moist tropical forests. *New Phytol.* **219**, 851–869 (2018). doi: [10.1111/nph.15027](https://doi.org/10.1111/nph.15027); pmid: [29451313](https://pubmed.ncbi.nlm.nih.gov/29451313/)
8. J. Carnicer *et al.*, Widespread crown condition decline, food web disruption, and amplified tree mortality with increased climate change-type drought. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1474–1478 (2011). doi: [10.1073/pnas.1010070108](https://doi.org/10.1073/pnas.1010070108); pmid: [21220333](https://pubmed.ncbi.nlm.nih.gov/21220333/)
9. C. Senf *et al.*, Canopy mortality has doubled in Europe's temperate forests over the last three decades. *Nat. Commun.* **9**, 4978 (2018). doi: [10.1038/s41467-018-07539-6](https://doi.org/10.1038/s41467-018-07539-6); pmid: [30478255](https://pubmed.ncbi.nlm.nih.gov/30478255/)
10. M. Fernández-Martínez *et al.*, Global trends in carbon sinks and their relationships with CO₂ and temperature. *Nat. Clim. Change* **9**, 73–79 (2019). doi: [10.1038/s41558-018-0367-7](https://doi.org/10.1038/s41558-018-0367-7)
11. P. Friedlingstein *et al.*, Global carbon budget 2019. *Earth Syst. Sci. Data* **11**, 1783–1838 (2019). doi: [10.5194/essd-11-1783-2019](https://doi.org/10.5194/essd-11-1783-2019)
12. P. Ciais *et al.*, Five decades of northern land carbon uptake revealed by the interhemispheric CO₂ gradient. *Nature* **568**, 221–225 (2019). doi: [10.1038/s41586-019-1078-6](https://doi.org/10.1038/s41586-019-1078-6); pmid: [30944480](https://pubmed.ncbi.nlm.nih.gov/30944480/)
13. J. M. Chen *et al.*, Vegetation structural change since 1981 significantly enhanced the terrestrial carbon sink. *Nat. Commun.* **10**, 4259 (2019). doi: [10.1038/s41467-019-12257-8](https://doi.org/10.1038/s41467-019-12257-8); pmid: [31534135](https://pubmed.ncbi.nlm.nih.gov/31534135/)
14. T. A. M. Pugh *et al.*, Role of forest regrowth in global carbon sink dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 4382–4387 (2019). doi: [10.1073/pnas.1810512116](https://doi.org/10.1073/pnas.1810512116); pmid: [30782807](https://pubmed.ncbi.nlm.nih.gov/30782807/)
15. J. Peñuelas *et al.*, Shifting from a fertilization-dominated to a warming-dominated period. *Nat. Ecol. Evol.* **1**, 1438–1445 (2017). doi: [10.1038/s41559-017-0274-8](https://doi.org/10.1038/s41559-017-0274-8); pmid: [29185529](https://pubmed.ncbi.nlm.nih.gov/29185529/)
16. W. Yuan *et al.*, Increased atmospheric vapor pressure deficit reduces global vegetation growth. *Sci. Adv.* **5**, eaax1396 (2019). doi: [10.1126/sciadv.aax1396](https://doi.org/10.1126/sciadv.aax1396); pmid: [31453338](https://pubmed.ncbi.nlm.nih.gov/31453338/)
17. K. F. Raffa *et al.*, Cross-scale drivers of natural disturbances prone to anthropogenic amplification: The dynamics of bark beetle eruptions. *Bioscience* **58**, 501–517 (2008). doi: [10.1641/B580607](https://doi.org/10.1641/B580607)
18. M. K. Tippett, C. Lepore, J. E. Cohen, More tornadoes in the most extreme U.S. tornado outbreaks. *Science* **354**, 1419–1423 (2016). doi: [10.1126/science.aah7393](https://doi.org/10.1126/science.aah7393); pmid: [27934705](https://pubmed.ncbi.nlm.nih.gov/27934705/)
19. G. R. van der Werf *et al.*, Global fire emissions estimates during 1997–2016. *Earth Syst. Sci. Data* **9**, 697–720 (2017). doi: [10.5194/essd-9-697-2017](https://doi.org/10.5194/essd-9-697-2017)
20. A. Sommerfeld *et al.*, Patterns and drivers of recent disturbances across the temperate forest biome. *Nat. Commun.* **9**, 4355 (2018). doi: [10.1038/s41467-018-06788-9](https://doi.org/10.1038/s41467-018-06788-9); pmid: [30341309](https://pubmed.ncbi.nlm.nih.gov/30341309/)
21. R. A. Houghton, A. A. Nassikas, Global and regional fluxes of carbon from land use and land cover change 1850–2015. *Global Biogeochem. Cycles* **31**, 456–472 (2017). doi: [10.1002/2016GB005546](https://doi.org/10.1002/2016GB005546)
22. J. A. Foley *et al.*, Global consequences of land use. *Science* **309**, 570–574 (2005). doi: [10.1126/science.1111772](https://doi.org/10.1126/science.1111772); pmid: [16040698](https://pubmed.ncbi.nlm.nih.gov/16040698/)
23. M. P. Perring *et al.*, Global environmental change effects on ecosystems: The importance of land-use legacies. *Glob. Change Biol.* **22**, 1361–1371 (2016). doi: [10.1111/gcb.13146](https://doi.org/10.1111/gcb.13146); pmid: [26546049](https://pubmed.ncbi.nlm.nih.gov/26546049/)
24. H. Hartmann *et al.*, Monitoring global tree mortality patterns and trends. Report from the VW symposium 'Crossing scales and disciplines to identify global trends of tree mortality as indicators of forest health'. *New Phytol.* **217**, 984–987 (2018). doi: [10.1111/nph.14988](https://doi.org/10.1111/nph.14988); pmid: [29334597](https://pubmed.ncbi.nlm.nih.gov/29334597/)
25. T. A. M. Pugh, A. Arneeth, M. Kautz, B. Poulter, B. Smith, Important role of forest disturbances in the global biomass turnover and carbon sinks. *Nat. Geosci.* **12**, 730–735 (2019). doi: [10.1038/s41561-019-0427-2](https://doi.org/10.1038/s41561-019-0427-2); pmid: [31478009](https://pubmed.ncbi.nlm.nih.gov/31478009/)
26. A. E. L. Stovall, H. Shugart, X. Yang, Tree height explains mortality risk during an intense drought. *Nat. Commun.* **10**, 4385 (2019). doi: [10.1038/s41467-019-12380-6](https://doi.org/10.1038/s41467-019-12380-6); pmid: [31558795](https://pubmed.ncbi.nlm.nih.gov/31558795/)
27. G. C. Hurtt *et al.*, Harmonization of Global Land-Use Change and Management for the Period 850–2100 (LUH2) for CMIP6. *Geosci. Model Dev. Discuss.* 10.5194/gmd-2019-360 (2020).
28. M. C. Hansen *et al.*, High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013). doi: [10.1126/science.1244693](https://doi.org/10.1126/science.1244693); pmid: [24233722](https://pubmed.ncbi.nlm.nih.gov/24233722/)
29. P. G. Curtis, C. M. Slay, N. L. Harris, A. Tyukavina, M. C. Hansen, Classifying drivers of global forest loss. *Science* **361**, 1108–1111 (2018). doi: [10.1126/science.aau3445](https://doi.org/10.1126/science.aau3445); pmid: [30213911](https://pubmed.ncbi.nlm.nih.gov/30213911/)
30. S. Froliking *et al.*, Forest disturbance and recovery: A general review in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy structure. *J. Geophys. Res.* **114**, G00E02 (2009). doi: [10.1029/2008JG000911](https://doi.org/10.1029/2008JG000911)
31. M. Kautz, A. J. H. Meddens, R. J. Hall, A. Arneeth, Biotic disturbances in Northern Hemisphere forests—A synthesis of recent data, uncertainties and implications for forest

- monitoring and modelling. *Glob. Ecol. Biogeogr.* **26**, 533–552 (2017). doi: [10.1111/geb.12558](https://doi.org/10.1111/geb.12558)
32. B. Poulter *et al.*, The Global Forest Age Dataset and its Uncertainties (GFADv1.1), NASA National Aeronautics and Space Administration, PANGAEA (2019). doi: [10.1594/PANGAEA.897392](https://doi.org/10.1594/PANGAEA.897392)
 33. S. L. LaDeau, J. S. Clark, Rising CO₂ levels and the fecundity of forest trees. *Science* **292**, 95–98 (2001). doi: [10.1126/science.1057547](https://doi.org/10.1126/science.1057547); pmid: [11292871](https://pubmed.ncbi.nlm.nih.gov/11292871/)
 34. J. E. Mohan, J. S. Clark, W. H. Schlesinger, Long-term CO₂ enrichment of a forest ecosystem: Implications for forest regeneration and succession. *Ecol. Appl.* **17**, 1198–1212 (2007). doi: [10.1890/05-1690](https://doi.org/10.1890/05-1690); pmid: [17555228](https://pubmed.ncbi.nlm.nih.gov/17555228/)
 35. L. G. Perry, P. B. Shafroth, D. M. Blumenthal, J. A. Morgan, D. R. LeCain, Elevated CO₂ does not offset greater water stress predicted under climate change for native and exotic riparian plants. *New Phytol.* **197**, 532–543 (2013). doi: [10.1111/nph.12030](https://doi.org/10.1111/nph.12030); pmid: [23171384](https://pubmed.ncbi.nlm.nih.gov/23171384/)
 36. N. Saintilan, K. Rogers, Woody plant encroachment of grasslands: A comparison of terrestrial and wetland settings. *New Phytol.* **205**, 1062–1070 (2015). doi: [10.1111/nph.13147](https://doi.org/10.1111/nph.13147); pmid: [25729806](https://pubmed.ncbi.nlm.nih.gov/25729806/)
 37. S. M. McMahon, G. G. Parker, D. R. Miller, Evidence for a recent increase in forest growth. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 3611–3615 (2010). doi: [10.1073/pnas.0912376107](https://doi.org/10.1073/pnas.0912376107); pmid: [21033710](https://pubmed.ncbi.nlm.nih.gov/21033710/)
 38. J. J. Camarero, A. Gazol, J. D. Galván, G. Sangüesa-Barreda, E. Gutiérrez, Disparate effects of global-change drivers on mountain conifer forests: Warming-induced growth enhancement in young trees vs. CO₂ fertilization in old trees from wet sites. *Glob. Change Biol.* **21**, 738–749 (2015). doi: [10.1111/gcb.12787](https://doi.org/10.1111/gcb.12787); pmid: [25362899](https://pubmed.ncbi.nlm.nih.gov/25362899/)
 39. S. L. Voecker, R. M. Muzika, R. P. Guyette, M. C. Stambaugh, Historical CO₂ growth enhancement declines with age in *Quercus* and *Pinus*. *Ecol. Monogr.* **76**, 549–564 (2006). doi: [10.1890/0012-9615\(2006\)076\[0549:HCGEDW\]2.0.CO;2](https://doi.org/10.1890/0012-9615(2006)076[0549:HCGEDW]2.0.CO;2)
 40. J. Peñuelas, J. G. Canadell, R. Ogaya, Increased water-use efficiency during the 20th century did not translate into enhanced tree growth. *Glob. Ecol. Biogeogr.* **20**, 597–608 (2011). doi: [10.1111/j.1466-8238.2010.00608.x](https://doi.org/10.1111/j.1466-8238.2010.00608.x)
 41. P. van der Sleen *et al.*, No growth stimulation of tropical trees by 150 years of CO₂ fertilization but water-use efficiency increased. *Nat. Geosci.* **8**, 24–28 (2015). doi: [10.1038/ngeo2313](https://doi.org/10.1038/ngeo2313)
 42. M. P. Girardin *et al.*, No growth stimulation of Canada's boreal forest under half-century of combined warming and CO₂ fertilization. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E8406–E8414 (2016). doi: [10.1073/pnas.1610156113](https://doi.org/10.1073/pnas.1610156113); pmid: [27956624](https://pubmed.ncbi.nlm.nih.gov/27956624/)
 43. R. J. W. Brienen *et al.*, Long-term decline of the Amazon carbon sink. *Nature* **519**, 344–348 (2015). doi: [10.1038/nature14283](https://doi.org/10.1038/nature14283); pmid: [25788097](https://pubmed.ncbi.nlm.nih.gov/25788097/)
 44. Z. Gedalof, A. Berg, Tree ring evidence for limited direct CO₂ fertilization of forests over the 20th century. *Global Biogeochem. Cycles* **24**, GB3027 (2010). doi: [10.1029/2009GB003699](https://doi.org/10.1029/2009GB003699)
 45. A. P. Walker *et al.*, Decadal biomass increment in early secondary succession woody ecosystems is increased by CO₂ enrichment. *Nat. Commun.* **10**, 454 (2019). doi: [10.1038/s41467-019-08348-1](https://doi.org/10.1038/s41467-019-08348-1); pmid: [30765702](https://pubmed.ncbi.nlm.nih.gov/30765702/)
 46. M. K. F. Bader *et al.*, Central European hardwood trees in a high-CO₂ future: Synthesis of an 8-year forest canopy CO₂ enrichment project. *J. Ecol.* **101**, 1509–1519 (2013). doi: [10.1111/1365-2745.12149](https://doi.org/10.1111/1365-2745.12149)
 47. D. S. Ellsworth *et al.*, Elevated CO₂ does not increase eucalypt forest productivity on a low-phosphorus soil. *Nat. Clim. Change* **7**, 279–282 (2017). doi: [10.1038/nclimate3235](https://doi.org/10.1038/nclimate3235)
 48. R. J. Norby *et al.*, Net primary productivity of a CO₂-enriched deciduous forest and the implications for carbon storage. *Ecol. Appl.* **12**, 1261–1266 (2002). doi: [10.1890/1051-0761\(2002\)012\[1261:NPP0AC\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2002)012[1261:NPP0AC]2.0.CO;2)
 49. N. G. McDowell, B. J. Bond, L. T. Dickman, M. G. Ryan, D. Whitehead, "Relationships between tree height and carbon isotope discrimination" in *Size- and Age-Related Changes in Tree Structure and Function*, F. C. Meinzer, B. Lachenbruch, T. E. Dawson, Eds. (Springer, 2011), pp. 255–286.
 50. H. Duan *et al.*, Elevated [CO₂] does not ameliorate the negative effects of elevated temperature on drought-induced mortality in *Eucalyptus radiata* seedlings. *Plant Cell Environ.* **37**, 1598–1613 (2014). doi: [10.1111/pce.12260](https://doi.org/10.1111/pce.12260); pmid: [24372529](https://pubmed.ncbi.nlm.nih.gov/24372529/)
 51. C. D. Allen, D. D. Breshears, N. G. McDowell, On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. *Ecosphere* **6**, 129 (2015). doi: [10.1890/ES15-00203.1](https://doi.org/10.1890/ES15-00203.1)
 52. C. Körner, A matter of tree longevity. *Science* **355**, 130–131 (2017). doi: [10.1126/science.aal2449](https://doi.org/10.1126/science.aal2449); pmid: [28082545](https://pubmed.ncbi.nlm.nih.gov/28082545/)
 53. U. Büntgen *et al.*, Limited capacity of tree growth to mitigate the global greenhouse effect under predicted warming. *Nat. Commun.* **10**, 2171 (2019). doi: [10.1038/s41467-019-10174-4](https://doi.org/10.1038/s41467-019-10174-4); pmid: [31092831](https://pubmed.ncbi.nlm.nih.gov/31092831/)
 54. A. C. Bennett, N. G. McDowell, C. D. Allen, K. J. Anderson-Teixeira, Larger trees suffer most during drought in forests worldwide. *Nat. Plants* **1**, 15139 (2015). doi: [10.1038/nplants.2015.139](https://doi.org/10.1038/nplants.2015.139); pmid: [27251391](https://pubmed.ncbi.nlm.nih.gov/27251391/)
 55. K. Yu *et al.*, Pervasive decreases in living vegetation carbon turnover time across forest climate zones. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 24662–24667 (2019). doi: [10.1073/pnas.1821387116](https://doi.org/10.1073/pnas.1821387116); pmid: [31740604](https://pubmed.ncbi.nlm.nih.gov/31740604/)
 56. H. Pretzsch, P. Biber, G. Schütze, J. Kemmerer, E. Uhl, Wood density reduced while wood volume growth accelerated in Central European forests since 1870. *For. Ecol. Manage.* **429**, 589–616 (2018). doi: [10.1016/j.foreco.2018.07.045](https://doi.org/10.1016/j.foreco.2018.07.045)
 57. K. E. Trenberth *et al.*, Global warming and changes in drought. *Nat. Clim. Change* **4**, 17–22 (2014). doi: [10.1038/nclimate2067](https://doi.org/10.1038/nclimate2067)
 58. A. Park Williams *et al.*, Temperature as a potent driver of regional forest drought stress and tree mortality. *Nat. Clim. Change* **3**, 292–297 (2013). doi: [10.1038/nclimate1693](https://doi.org/10.1038/nclimate1693)
 59. A. J. Tepley, J. R. Thompson, H. E. Epstein, K. J. Anderson-Teixeira, Vulnerability to forest loss through altered postfire recovery dynamics in a warming climate in the Klamath Mountains. *Glob. Change Biol.* **23**, 4117–4132 (2017). doi: [10.1111/gcb.13704](https://doi.org/10.1111/gcb.13704); pmid: [28447370](https://pubmed.ncbi.nlm.nih.gov/28447370/)
 60. J. M. Serra-Diaz *et al.*, Disequilibrium of fire-prone forests sets the stage for a rapid decline in conifer dominance during the 21st century. *Sci. Rep.* **8**, 6749 (2018). doi: [10.1038/s41598-018-24642-2](https://doi.org/10.1038/s41598-018-24642-2); pmid: [29712940](https://pubmed.ncbi.nlm.nih.gov/29712940/)
 61. M. Uriarte, J. R. Lasky, V. K. Boukili, R. L. Chazdon, A trait-mediated, neighbourhood approach to quantify climate impacts on successional dynamics of tropical rainforests. *Funct. Ecol.* **30**, 157–167 (2016). doi: [10.1111/1365-2435.12576](https://doi.org/10.1111/1365-2435.12576)
 62. N. G. McDowell, C. D. Allen, Darcy's law predicts widespread forest mortality under climate warming. *Nat. Clim. Change* **5**, 669–672 (2015). doi: [10.1038/nclimate2641](https://doi.org/10.1038/nclimate2641)
 63. J. Liu *et al.*, Contrasting carbon cycle responses of the tropical continents to the 2015–2016 El Niño. *Science* **358**, eaam5690 (2017). doi: [10.1126/science.aam5690](https://doi.org/10.1126/science.aam5690); pmid: [29026011](https://pubmed.ncbi.nlm.nih.gov/29026011/)
 64. S. J. Wright, O. Calderón, Seasonal, El Niño and longer term changes in flower and seed production in a moist tropical forest. *Ecol. Lett.* **9**, 35–44 (2006). pmid: [16958866](https://pubmed.ncbi.nlm.nih.gov/16958866/)
 65. T. F. Keenan, W. J. Riley, Greening of the land surface in the world's cold regions consistent with recent warming. *Nat. Clim. Change* **8**, 825–828 (2018). doi: [10.1038/s41558-018-0258-y](https://doi.org/10.1038/s41558-018-0258-y); pmid: [30319714](https://pubmed.ncbi.nlm.nih.gov/30319714/)
 66. J. R. Forrest, Plant–pollinator interactions and phenological change: What can we learn about climate impacts from experiments and observations? *Oikos* **124**, 4–13 (2015). doi: [10.1111/oik.01386](https://doi.org/10.1111/oik.01386)
 67. J. L. Walck, S. N. Hidayati, K. W. Dixon, K. Thompson, P. Poschlod, Climate change and plant regeneration from seed. *Glob. Change Biol.* **17**, 2145–2161 (2011). doi: [10.1111/j.1365-2486.2010.02368.x](https://doi.org/10.1111/j.1365-2486.2010.02368.x)
 68. J. Memmott, P. G. Craze, N. M. Waser, M. V. Price, Global warming and the disruption of plant–pollinator interactions. *Ecol. Lett.* **10**, 710–717 (2007). doi: [10.1111/j.1461-0248.2007.01061.x](https://doi.org/10.1111/j.1461-0248.2007.01061.x); pmid: [17594426](https://pubmed.ncbi.nlm.nih.gov/17594426/)
 69. L. M. Kueppers *et al.*, Warming and provenance limit tree recruitment across and beyond the elevation range of subalpine forest. *Glob. Change Biol.* **23**, 2383–2395 (2017). doi: [10.1111/gcb.13561](https://doi.org/10.1111/gcb.13561); pmid: [27976819](https://pubmed.ncbi.nlm.nih.gov/27976819/)
 70. W. D. Hansen, M. G. Turner, Origins of abrupt change? Postfire subalpine conifer regeneration declines nonlinearly with warming and drying. *Ecol. Monogr.* **89**, e01340 (2019). doi: [10.1002/ecm.1340](https://doi.org/10.1002/ecm.1340)
 71. H. D. Adams *et al.*, Temperature sensitivity of drought-induced tree mortality portends increased regional die-off under global-change-type drought. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 7063–7066 (2009). doi: [10.1073/pnas.0901438106](https://doi.org/10.1073/pnas.0901438106); pmid: [19365070](https://pubmed.ncbi.nlm.nih.gov/19365070/)
 72. R. A. Hember *et al.*, Accelerating regrowth of temperate-maritime forests due to environmental change. *Glob. Clim. Change* **18**, 2026–2040 (2012). doi: [10.1111/j.1365-2486.2012.02669.x](https://doi.org/10.1111/j.1365-2486.2012.02669.x)
 73. H. D. Adams *et al.*, A multi-species synthesis of physiological mechanisms in drought-induced tree mortality. *Nat. Ecol. Evol.* **1**, 1285–1291 (2017). doi: [10.1038/s41559-017-0248-x](https://doi.org/10.1038/s41559-017-0248-x); pmid: [29046541](https://pubmed.ncbi.nlm.nih.gov/29046541/)
 74. M. L. Gaylor *et al.*, Drought predisposes piñon-juniper woodlands to insect attacks and mortality. *New Phytol.* **198**, 567–578 (2013). doi: [10.1111/nph.12174](https://doi.org/10.1111/nph.12174); pmid: [23421561](https://pubmed.ncbi.nlm.nih.gov/23421561/)
 75. N. G. McDowell *et al.*, Multi-scale predictions of massive conifer mortality due to chronic temperature rise. *Nat. Clim. Change* **6**, 295–300 (2016). doi: [10.1038/nclimate2873](https://doi.org/10.1038/nclimate2873)
 76. B. M. J. Engelbrecht, T. A. Kursar, Comparative drought-resistance of seedlings of 28 species of co-occurring tropical woody plants. *Oecologia* **136**, 383–393 (2003). doi: [10.1007/s00442-003-1290-8](https://doi.org/10.1007/s00442-003-1290-8); pmid: [12811534](https://pubmed.ncbi.nlm.nih.gov/12811534/)
 77. C. E. Doughty *et al.*, Drought impact on forest carbon dynamics and fluxes in Amazonia. *Nature* **519**, 78–82 (2015). doi: [10.1038/nature14213](https://doi.org/10.1038/nature14213); pmid: [25739631](https://pubmed.ncbi.nlm.nih.gov/25739631/)
 78. K. H. Erb *et al.*, Land management: Data availability and process understanding for global change studies. *Glob. Change Biol.* **23**, 512–533 (2017). doi: [10.1111/gcb.13443](https://doi.org/10.1111/gcb.13443); pmid: [27447350](https://pubmed.ncbi.nlm.nih.gov/27447350/)
 79. P. A. Martin, A. C. Newton, M. Pfeifer, M. Khoo, J. M. Bullock, Impacts of tropical selective logging on carbon storage and tree species richness: A meta-analysis. *For. Ecol. Manage.* **356**, 224–233 (2015). doi: [10.1016/j.foreco.2015.07.010](https://doi.org/10.1016/j.foreco.2015.07.010)
 80. A. Chaudhary, Z. Burivalova, L. P. Koh, S. Hellweg, Impact of forest management on species richness: Global meta-analysis and economic trade-offs. *Sci. Rep.* **6**, 23954 (2016). doi: [10.1038/srep23954](https://doi.org/10.1038/srep23954); pmid: [27040604](https://pubmed.ncbi.nlm.nih.gov/27040604/)
 81. M. J. Duveneck, J. R. Thompson, E. J. Gustafson, Y. Liang, A. M. G. de Bruijn, Recovery dynamics and climate change effects to future New England forests. *Lands. Ecol.* **32**, 1385–1397 (2017). doi: [10.1007/s10980-016-0415-5](https://doi.org/10.1007/s10980-016-0415-5)
 82. D. Thom, W. Rammer, R. Garstenauer, R. Seidl, Legacies of past land use have a stronger effect on forest carbon exchange than future climate change in a temperate forest landscape. *Biogeosciences* **15**, 5699–5713 (2018). doi: [10.5194/bg-15-5699-2018](https://doi.org/10.5194/bg-15-5699-2018)
 83. T. Vilen *et al.*, Reconstructed forest age structure in Europe 1950–2010. *For. Ecol. Manage.* **286**, 203–218 (2012). doi: [10.1016/j.foreco.2012.08.048](https://doi.org/10.1016/j.foreco.2012.08.048)
 84. W. M. Jolly *et al.*, Climate-induced variations in global wildfire danger from 1979 to 2013. *Nat. Commun.* **6**, 7537 (2015). doi: [10.1038/ncomms8537](https://doi.org/10.1038/ncomms8537); pmid: [26172867](https://pubmed.ncbi.nlm.nih.gov/26172867/)
 85. N. Andela *et al.*, A human-driven decline in global burned area. *Science* **356**, 1356–1362 (2017). doi: [10.1126/science.aal4108](https://doi.org/10.1126/science.aal4108); pmid: [28663495](https://pubmed.ncbi.nlm.nih.gov/28663495/)
 86. J. T. Abatzoglou, A. P. Williams, Impact of anthropogenic climate change on wildfire across western US forests. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 11770–11775 (2016). doi: [10.1073/pnas.1607111113](https://doi.org/10.1073/pnas.1607111113); pmid: [27791053](https://pubmed.ncbi.nlm.nih.gov/27791053/)
 87. A. L. Westerling, M. G. Turner, E. A. H. Smithwick, W. H. Romme, M. K. Ryan, Continued warming could transform Greater Yellowstone fire regimes by mid-21st century. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13165–13170 (2011). doi: [10.1073/pnas.1110199108](https://doi.org/10.1073/pnas.1110199108); pmid: [21788495](https://pubmed.ncbi.nlm.nih.gov/21788495/)
 88. D. Bowman, G. J. Williamson, L. D. Prior, B. P. Murphy, The relative importance of intrinsic and extrinsic factors in the decline of obligate seeder forests. *Glob. Ecol. Biogeogr.* **25**, 1166–1172 (2016). doi: [10.1111/geb.12484](https://doi.org/10.1111/geb.12484)
 89. J. F. Johnstone *et al.*, Changing disturbance regimes, ecological memory, and forest resilience. *Front. Ecol. Environ.* **14**, 369–378 (2016). doi: [10.1002/fee.1311](https://doi.org/10.1002/fee.1311)
 90. M. G. Turner, K. H. Brazilians, W. D. Hansen, B. J. Harvey, Short-interval severe fire erodes the resilience of subalpine lodgepole pine forests. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 11319–11328 (2019). doi: [10.1073/pnas.1902841116](https://doi.org/10.1073/pnas.1902841116); pmid: [31110003](https://pubmed.ncbi.nlm.nih.gov/31110003/)
 91. T. Kitzberger *et al.*, Fire-vegetation feedbacks and alternative states: Common mechanisms of temperate forest vulnerability to fire in southern South America and New Zealand. *N. Z. J. Bot.* **54**, 247–272 (2016). doi: [10.1080/0028825X.2016.1151903](https://doi.org/10.1080/0028825X.2016.1151903)
 92. A. E. Lugo, Visible and invisible effects of hurricanes on forest ecosystems: An international review. *Austral Ecol.* **33**, 368–398 (2008). doi: [10.1111/j.1442-9993.2008.01894.x](https://doi.org/10.1111/j.1442-9993.2008.01894.x)
 93. K. Balaguru, G. R. Foltz, L. R. Leung, Increasing magnitude of hurricane rapid intensification in the central and eastern tropical Atlantic. *Geophys. Res. Lett.* **45**, 4238–4247 (2018). doi: [10.1029/2018GL077597](https://doi.org/10.1029/2018GL077597)
 94. N. S. Diefenbaugh, M. Scherer, R. J. Trapp, Robust increases in severe thunderstorm environments in response to greenhouse forcing. *Proc. Natl. Acad. Sci. U.S.A.* **110**,

- 16361–16366 (2013). doi: [10.1073/pnas.1307758110](https://doi.org/10.1073/pnas.1307758110); pmid: [24062439](https://pubmed.ncbi.nlm.nih.gov/24062439/)
95. M. Uriarte, C. D. Canham, J. Thompson, J. K. Zimmerman, A neighborhood analysis of tree growth and survival in a hurricane-driven tropical forest. *Ecol. Monogr.* **74**, 591–614 (2004). doi: [10.1890/03-4031](https://doi.org/10.1890/03-4031)
 96. B. Gardiner, P. Berry, B. Mouli, Review: Wind impacts on plant growth, mechanics and damage. *Plant Sci.* **245**, 94–118 (2016). doi: [10.1016/j.plantsci.2016.01.006](https://doi.org/10.1016/j.plantsci.2016.01.006); pmid: [26940495](https://pubmed.ncbi.nlm.nih.gov/26940495/)
 97. M. Uriarte *et al.*, Natural disturbances and human land use as determinants of tropical forest dynamics: Results from a forest simulator. *Ecol. Monogr.* **79**, 423–443 (2009). doi: [10.1890/08-0707.1](https://doi.org/10.1890/08-0707.1)
 98. L. S. Comita *et al.*, Abiotic and biotic drivers of seedling survival in a hurricane-impacted tropical forest. *J. Ecol.* **97**, 1346–1359 (2009). doi: [10.1111/j.1365-2745.2009.01551.x](https://doi.org/10.1111/j.1365-2745.2009.01551.x)
 99. M. Uriarte *et al.*, Multidimensional trade-offs in species responses to disturbance: Implications for diversity in a subtropical forest. *Ecology* **93**, 191–205 (2012). doi: [10.1890/10-2422.1](https://doi.org/10.1890/10-2422.1); pmid: [22486099](https://pubmed.ncbi.nlm.nih.gov/22486099/)
 100. D. F. B. Flynn *et al.*, Hurricane disturbance alters secondary forest recovery in Puerto Rico. *Biotropica* **42**, 149–157 (2010). doi: [10.1111/j.1744-7429.2009.00581.x](https://doi.org/10.1111/j.1744-7429.2009.00581.x)
 101. S. P. Yanoviak *et al.*, Lightning is a major cause of large tree mortality in a lowland neotropical forest. *New Phytol.* **225**, 1936–1944 (2020). doi: [10.1111/nph.16260](https://doi.org/10.1111/nph.16260); pmid: [31610011](https://pubmed.ncbi.nlm.nih.gov/31610011/)
 102. R. Seidl, M. J. Schelhaas, W. Rammer, P. J. Verkerk, Increasing forest disturbances in Europe and their impact on carbon storage. *Nat. Clim. Change* **4**, 806–810 (2014). doi: [10.1038/nclimate2318](https://doi.org/10.1038/nclimate2318); pmid: [25377744](https://pubmed.ncbi.nlm.nih.gov/25377744/)
 103. K. F. Raffa *et al.*, “Responses of tree-killing bark beetles to a changing climate” in *Climate Change and Insect Pests*, C. Björkman, P. Niemelä, Eds., vol. 7 of *CABI Climate Change Series* (CABI International, 2015), pp. 173–201.
 104. J. S. Schurman *et al.*, Large-scale disturbance legacies and the climate sensitivity of primary *Picea abies* forests. *Glob. Change Biol.* **24**, 2169–2181 (2018). doi: [10.1111/gcb.14041](https://doi.org/10.1111/gcb.14041); pmid: [29322582](https://pubmed.ncbi.nlm.nih.gov/29322582/)
 105. D. W. Rosenberger, R. C. Venette, B. H. Aukema, Development of an aggressive bark beetle on novel hosts: Implications for outbreaks in an invaded range. *J. Appl. Ecol.* **55**, 1526–1537 (2018). doi: [10.1111/1365-2664.13064](https://doi.org/10.1111/1365-2664.13064)
 106. R. Seidl *et al.*, Invasive alien pests threaten the carbon stored in Europe’s forests. *Nat. Commun.* **9**, 1626 (2018). doi: [10.1038/s41467-018-04096-w](https://doi.org/10.1038/s41467-018-04096-w); pmid: [29691396](https://pubmed.ncbi.nlm.nih.gov/29691396/)
 107. D. M. Johnson *et al.*, Climatic warming disrupts recurrent Alpine insect outbreaks. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 20576–20581 (2010). doi: [10.1073/pnas.1010270107](https://doi.org/10.1073/pnas.1010270107); pmid: [21059922](https://pubmed.ncbi.nlm.nih.gov/21059922/)
 108. M. di Porcia e Brugnera *et al.*, Modeling the impact of liana infestation on the demography and carbon cycle of tropical forests. *Glob. Change Biol.* **25**, 3767–3780 (2019). doi: [10.1111/gcb.14769](https://doi.org/10.1111/gcb.14769); pmid: [31310429](https://pubmed.ncbi.nlm.nih.gov/31310429/)
 109. D. S. Pureswaran, R. Johns, S. B. Heard, D. Quiring, Paradigms in eastern spruce budworm (Lepidoptera: Tortricidae) population ecology: a century of debate. *Environ. Entomol.* **45**, 1333–1342 (2016). doi: [10.1093/ee/nwv103](https://doi.org/10.1093/ee/nwv103); pmid: [28028079](https://pubmed.ncbi.nlm.nih.gov/28028079/)
 110. G. V. J. Virgin, D. A. MacLean, Five decades of balsam fir stand development after spruce budworm-related mortality. *For. Ecol. Manage.* **400**, 129–138 (2017). doi: [10.1016/j.foreco.2017.05.057](https://doi.org/10.1016/j.foreco.2017.05.057)
 111. M. Macek *et al.*, Life and death of *Picea abies* after bark-beetle outbreak: Ecological processes driving seedling recruitment. *Ecol. Appl.* **27**, 156–167 (2017). doi: [10.1002/eap.1429](https://doi.org/10.1002/eap.1429); pmid: [28052495](https://pubmed.ncbi.nlm.nih.gov/28052495/)
 112. W. F. Laurance, P. Delamônica, S. G. Laurance, H. L. Vasconcelos, T. E. Lovejoy, Rainforest fragmentation kills big trees. *Nature* **404**, 836 (2000). doi: [10.1038/35009032](https://doi.org/10.1038/35009032); pmid: [10786782](https://pubmed.ncbi.nlm.nih.gov/10786782/)
 113. D. B. Lindenmayer, W. F. Laurance, J. F. Franklin, Global decline in large old trees. *Science* **338**, 1305–1306 (2012). doi: [10.1126/science.1231070](https://doi.org/10.1126/science.1231070); pmid: [23224548](https://pubmed.ncbi.nlm.nih.gov/23224548/)
 114. C. J. Peterson, Catastrophic wind damage to North American forests and the potential impact of climate change. *Sci. Total Environ.* **262**, 287–311 (2000). doi: [10.1016/S0048-9697\(00\)00529-5](https://doi.org/10.1016/S0048-9697(00)00529-5); pmid: [11087033](https://pubmed.ncbi.nlm.nih.gov/11087033/)
 115. C. D. Canham, M. J. Papaik, E. F. Latty, Interspecific variation in susceptibility to windthrow as a function of tree size and storm severity for northern temperate tree species. *Can. J. For. Res.* **31**, 1–10 (2001). doi: [10.1139/x00-124](https://doi.org/10.1139/x00-124)
 116. J. Boone, B. H. Aukema, J. Bohlmann, A. L. Carroll, K. F. Raffa, Efficacy of tree defense physiology varies with bark beetle population density: A basis for positive feedback in eruptive species. *Can. J. For. Res.* **41**, 1174–1188 (2011). doi: [10.1139/x11-041](https://doi.org/10.1139/x11-041)
 117. N. G. McDowell *et al.*, Predicting chronic climate-driven disturbances and their mitigation. *Trends Ecol. Evol.* **33**, 15–27 (2018). doi: [10.1016/j.tree.2017.10.002](https://doi.org/10.1016/j.tree.2017.10.002); pmid: [29146414](https://pubmed.ncbi.nlm.nih.gov/29146414/)
 118. D. J. Johnson *et al.*, Climate sensitive size-dependent survival in tropical trees. *Nat. Ecol. Evol.* **2**, 1436–1442 (2018). doi: [10.1038/s41559-018-0626-z](https://doi.org/10.1038/s41559-018-0626-z); pmid: [30104751](https://pubmed.ncbi.nlm.nih.gov/30104751/)
 119. H. Pretzsch, P. Biber, G. Schütze, E. Uhl, T. Rötzer, Forest stand growth dynamics in Central Europe have accelerated since 1870. *Nat. Commun.* **5**, 4967 (2014). doi: [10.1038/ncomms5967](https://doi.org/10.1038/ncomms5967); pmid: [25216297](https://pubmed.ncbi.nlm.nih.gov/25216297/)
 120. A. B. Nicotra *et al.*, Plant phenotypic plasticity in a changing climate. *Trends Plant Sci.* **15**, 684–692 (2010). doi: [10.1016/j.jplants.2010.09.008](https://doi.org/10.1016/j.jplants.2010.09.008); pmid: [20970368](https://pubmed.ncbi.nlm.nih.gov/20970368/)
 121. S. N. Aitken, S. Yeaman, J. A. Holliday, T. Wang, S. Curtis-McLane, Adaptation, migration or extirpation: Climate change outcomes for tree populations. *Evol. Appl.* **1**, 95–111 (2008). doi: [10.1111/j.1752-4571.2007.00013.x](https://doi.org/10.1111/j.1752-4571.2007.00013.x); pmid: [25567494](https://pubmed.ncbi.nlm.nih.gov/25567494/)
 122. A. L. Angert *et al.*, Do species’ traits predict recent shifts at expanding range edges? *Ecol. Lett.* **14**, 677–689 (2011). doi: [10.1111/j.1461-0248.2011.01620.x](https://doi.org/10.1111/j.1461-0248.2011.01620.x); pmid: [21535340](https://pubmed.ncbi.nlm.nih.gov/21535340/)
 123. K. M. Miller, B. J. McGill, Land use and life history limit migration capacity of eastern tree species. *Glob. Ecol. Biogeogr.* **27**, 57–67 (2018). doi: [10.1111/geb.12671](https://doi.org/10.1111/geb.12671)
 124. M. Slot, K. Kitajima, General patterns of acclimation of leaf respiration to elevated temperatures across biomes and plant types. *Oecologia* **177**, 885–900 (2015). doi: [10.1007/s00442-014-3159-4](https://doi.org/10.1007/s00442-014-3159-4); pmid: [25481817](https://pubmed.ncbi.nlm.nih.gov/25481817/)
 125. I. Rieu, D. Twell, N. Firon, Pollen development at high temperature: From acclimation to collapse. *Plant Physiol.* **173**, 1967–1976 (2017). doi: [10.1104/pp.16.01644](https://doi.org/10.1104/pp.16.01644); pmid: [28246296](https://pubmed.ncbi.nlm.nih.gov/28246296/)
 126. H. D. Adams *et al.*, Experimental drought and heat can delay phenological development and reduce foliar and shoot growth in semiarid trees. *Glob. Change Biol.* **21**, 4210–4220 (2015). doi: [10.1111/gcb.13030](https://doi.org/10.1111/gcb.13030); pmid: [26149972](https://pubmed.ncbi.nlm.nih.gov/26149972/)
 127. C. Grossiord *et al.*, Warming combined with more extreme precipitation regimes modifies the water sources used by trees. *New Phytol.* **213**, 584–596 (2017). doi: [10.1111/nph.14192](https://doi.org/10.1111/nph.14192); pmid: [27612306](https://pubmed.ncbi.nlm.nih.gov/27612306/)
 128. J. S. McLachlan, J. J. Hellmann, M. W. Schwartz, A framework for debate of assisted migration in an era of climate change. *Conserv. Biol.* **21**, 297–302 (2007). doi: [10.1111/j.1523-1739.2007.00676.x](https://doi.org/10.1111/j.1523-1739.2007.00676.x); pmid: [17391179](https://pubmed.ncbi.nlm.nih.gov/17391179/)
 129. F. Lloret, A. Escudero, J. M. Iriondo, J. Martínez-Vilalta, F. Valladares, Extreme climatic events and vegetation: The role of stabilizing processes. *Glob. Change Biol.* **18**, 797–805 (2012). doi: [10.1111/j.1365-2486.2011.02624.x](https://doi.org/10.1111/j.1365-2486.2011.02624.x)
 130. K. B. Kemp, P. E. Higuera, P. Morgan, Fire legacies impact conifer regeneration across environmental gradients in the U.S. northern Rockies. *Landsc. Ecol.* **31**, 619–636 (2016). doi: [10.1038/nature15374](https://doi.org/10.1038/nature15374); pmid: [26466564](https://pubmed.ncbi.nlm.nih.gov/26466564/)
 131. D. Thom, W. Rammer, R. Seidl, The impact of future forest dynamics on climate: Interactive effects of changing vegetation and disturbance regimes. *Ecol. Monogr.* **87**, 665–684 (2017). doi: [10.1002/ecm.1272](https://doi.org/10.1002/ecm.1272); pmid: [29628526](https://pubmed.ncbi.nlm.nih.gov/29628526/)
 132. T. L. Powell *et al.*, Variation in hydroclimate sustains tropical forest biomass and promotes functional diversity. *New Phytol.* **219**, 932–946 (2018). doi: [10.1111/nph.15271](https://doi.org/10.1111/nph.15271); pmid: [29923303](https://pubmed.ncbi.nlm.nih.gov/29923303/)
 133. G. B. Bonan, Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science* **320**, 1444–1449 (2008). doi: [10.1126/science.1155121](https://doi.org/10.1126/science.1155121); pmid: [18556546](https://pubmed.ncbi.nlm.nih.gov/18556546/)
 134. S. Luyssaert *et al.*, Old-growth forests as global carbon sinks. *Nature* **455**, 213–215 (2008). doi: [10.1038/nature07276](https://doi.org/10.1038/nature07276); pmid: [18784722](https://pubmed.ncbi.nlm.nih.gov/18784722/)
 135. M. E. Harmon, Carbon sequestration in forests: Addressing the scale question. *J. For.* **99**, 24–29 (2001).
 136. J. E. Campbell *et al.*, Large historical growth in global terrestrial gross primary production. *Nature* **544**, 84–87 (2017). doi: [10.1038/nature22030](https://doi.org/10.1038/nature22030); pmid: [28382993](https://pubmed.ncbi.nlm.nih.gov/28382993/)
 137. J. T. Randerson *et al.*, The impact of boreal forest fire on climate warming. *Science* **314**, 1130–1132 (2006). doi: [10.1126/science.1132075](https://doi.org/10.1126/science.1132075); pmid: [17110574](https://pubmed.ncbi.nlm.nih.gov/17110574/)
 138. X. Lee *et al.*, Observed increase in local cooling effect of deforestation at higher latitudes. *Nature* **479**, 384–387 (2011). doi: [10.1038/nature10588](https://doi.org/10.1038/nature10588); pmid: [22094699](https://pubmed.ncbi.nlm.nih.gov/22094699/)
 139. N. Devaraju, G. Bala, A. Modak, Effects of large-scale deforestation on precipitation in the monsoon regions: Remote versus local effects. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3257–3262 (2015). doi: [10.1073/pnas.1423439112](https://doi.org/10.1073/pnas.1423439112); pmid: [25733889](https://pubmed.ncbi.nlm.nih.gov/25733889/)
 140. Q. Lejeune, E. L. Davin, B. P. Guilloid, S. I. Seneviratne, Influence of Amazonian deforestation on the future evolution of regional surface fluxes, circulation, surface temperature and precipitation. *Clim. Dyn.* **44**, 2769–2786 (2015). doi: [10.1007/s00382-014-2203-8](https://doi.org/10.1007/s00382-014-2203-8)
 141. Y. Le Page, D. Morton, B. Bond-Lamberty, J. M. C. Pereira, G. Hurtt, HESFIRE: A global fire model to explore the role of anthropogenic and weather drivers. *Biogeosciences* **12**, 887–903 (2015). doi: [10.5194/bg-12-887-2015](https://doi.org/10.5194/bg-12-887-2015)
 142. N. G. McDowell *et al.*, Evaluating theories of drought-induced vegetation mortality using a multimodel-experiment framework. *New Phytol.* **200**, 304–321 (2013). doi: [10.1111/nph.12465](https://doi.org/10.1111/nph.12465); pmid: [24004027](https://pubmed.ncbi.nlm.nih.gov/24004027/)
 143. D. Lawrence *et al.*, The Land Use Model Intercomparison Project (LUMP) contribution to CMIP6: Rationale and experimental design. *Geosci. Model Dev.* **9**, 2973–2998 (2016). doi: [10.5194/gmd-9-2973-2016](https://doi.org/10.5194/gmd-9-2973-2016)
 144. M. Simard, N. Pinto, J. B. Fisher, A. Baccini, Mapping forest canopy height globally with spaceborne lidar. *J. Geophys. Res. Biogeosci.* **116**, G04021 (2011).
 145. V. Lehsten, M. Mischorow, E. Lindström, D. Lehsten, H. Lischke, LPJ-GM 1.0: Simulating migration efficiently in a dynamic vegetation model. *Geosci. Model Dev.* **12**, 893–908 (2019). doi: [10.5194/gmd-12-893-2019](https://doi.org/10.5194/gmd-12-893-2019)
 146. B. Smith, I. C. Prentice, M. T. Sykes, Representation of vegetation dynamics in the modelling of terrestrial ecosystems: Comparing two contrasting approaches within European climate space. *Glob. Ecol. Biogeogr.* **10**, 621–637 (2001). doi: [10.1046/j.1466-822X.2001.00256.x](https://doi.org/10.1046/j.1466-822X.2001.00256.x)
 147. R. A. Fisher *et al.*, Taking off the training wheels: The properties of a dynamic vegetation model without climate envelopes, CLM4.5(ED). *Geosci. Model Dev.* **8**, 3593–3619 (2015). doi: [10.5194/gmd-8-3593-2015](https://doi.org/10.5194/gmd-8-3593-2015)
 148. Y. Chen *et al.*, Simulating damage for wind storms in the land surface model ORCHIDEE-CAN (revision 4262). *Geosci. Model Dev.* **11**, 771–791 (2018). doi: [10.5194/gmd-11-771-2018](https://doi.org/10.5194/gmd-11-771-2018)
 149. T. K. Henkel, J. Q. Chambers, D. A. Baker, Delayed tree mortality and Chinese tallow (*Triadica sebifera*) population explosion in a Louisiana bottomland hardwood forest following Hurricane Katrina. *For. Ecol. Manage.* **378**, 222–232 (2016). doi: [10.1016/j.foreco.2016.07.036](https://doi.org/10.1016/j.foreco.2016.07.036)
 150. J. P. Roccaforte *et al.*, Delayed tree mortality, bark beetle activity, and regeneration dynamics five years following the Wallow Fire, Arizona, USA: Assessing trajectories towards resiliency. *For. Ecol. Manage.* **428**, 20–26 (2018). doi: [10.1016/j.foreco.2018.06.012](https://doi.org/10.1016/j.foreco.2018.06.012)
 151. D. Kennedy *et al.*, Implementing plant hydraulics in the Community Land Model, version 5. *J. Adv. Model. Earth Syst.* **11**, 485–513 (2019). doi: [10.1029/2018MS001500](https://doi.org/10.1029/2018MS001500)
 152. P. J. Grubb, The maintenance of species-richness in plant communities: The importance of the regeneration niche. *Biol. Rev. Camb. Philos. Soc.* **52**, 107–145 (1977). doi: [10.1111/j.1469-185X.1977.tb01347.x](https://doi.org/10.1111/j.1469-185X.1977.tb01347.x)
 153. B. D. Amiro *et al.*, Ecosystem carbon dioxide fluxes after disturbance in forests of North America. *J. Geophys. Res. Biogeosci.* **115**, G00K02 (2010).
 154. J. A. Hicke *et al.*, Effects of biotic disturbances on forest carbon cycling in the United States and Canada. *Glob. Change Biol.* **18**, 7–34 (2012). doi: [10.1111/j.1365-2486.2011.02543.x](https://doi.org/10.1111/j.1365-2486.2011.02543.x)
 155. J. P. Grime, *Plant Strategies and Vegetation Processes* (Wiley, 1979).
 156. A. Jentsch, P. White, A theory of pulse dynamics and disturbance in ecology. *Ecology* **100**, e02734 (2019). doi: [10.1002/ecy.2734](https://doi.org/10.1002/ecy.2734); pmid: [31018013](https://pubmed.ncbi.nlm.nih.gov/31018013/)
 157. D. M. Olson *et al.*, Terrestrial ecoregions of the world: A new map of life on Earth. *Bioscience* **51**, 933–938 (2001). doi: [10.1641/0006-3568\(2001\)051\[0933:TEOTWA\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2)

ACKNOWLEDGMENTS

Funding: This manuscript was derived from the Department of Energy (DOE) Workshop “Disturbance and vegetation dynamics in Earth System Models” held March 2018 in Washington, D.C. Funding was provided by: the DOE’s Next Generation Ecosystem Experiment-Tropics (NGEE-Tropics) and Pacific Northwest National Lab’s (PNNL’s) LDRD program (N.G.M.), McIntire-Stennis MIN-17-095 (B.H.A.), the U.S. Geological Survey’s Ecosystems and Land Resources mission areas (C.D.A.), the Joint Fire Science Program (16-3-01-4) and the University of Wisconsin–Madison

Vilas Trust (M.G.T.), sabbatical fellowship support from sDiv (J.W.L.), the Synthesis Centre of iDiv (DFG FZT 118), DOE's NGEE-Tropics and the University of California's Laboratory Fees Research Program (grant LFR-18-542511 to C.X.), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant 758873, TreeMort, to

T.A.M.P.), paper number 43 of the Birmingham Institute of Forest Research, PNNL's LDRD program (B.B.-L.), sabbatical support from Stanford's Center for Advanced Study in the Behavioral Sciences and financial support from the Gordon and Betty Moore Foundation (R.B.J.), the NASA Carbon Monitoring System and NASA Interdisciplinary Science Programs (G.C.H. and B.P.). Oak

Ridge National Laboratory is operated by UT-Battelle, LLC, under contract DE-AC05-00OR22725 to the United States Department of Energy. **Competing interests:** The authors have no competing interests to declare.

10.1126/science.aaz9463

RESEARCH ARTICLE SUMMARY

IMMUNOLOGY

Granzyme A from cytotoxic lymphocytes cleaves GSDMB to trigger pyroptosis in target cells

Zhiwei Zhou*, Huabin He*, Kun Wang, Xuyan Shi, Yupeng Wang, Ya Su, Yao Wang, Da Li, Wang Liu, Yongliang Zhang, Lianjun Shen, Weidong Han, Lin Shen, Jingjin Ding, Feng Shao†

INTRODUCTION: In cellular immunity, cytotoxic T lymphocytes (CTLs) and natural killer (NK) cells use perforin to deliver serine protease granzymes into target cells to kill them. Gasdermins are pore-forming proteins that execute pyroptosis, a form of proinflammatory cell death. Gasdermin D (GSDMD) is cleaved by caspase-1/4/5/11 upon inflammasome activation, releasing the pore-forming domain for plasma membrane disruption. Gasdermin E (GSDME) is similarly cleaved by caspase-3, converting apoptosis to pyroptosis. The functional mechanism for other gasdermins is unknown.

RATIONALE: The view that granzymes induce target-cell apoptosis was proposed two decades ago, when apoptosis was thought to be the

dominant form of programmed cell death and assays to ascertain apoptosis were insufficiently accurate. Furthermore, granzyme cytotoxicity was only assessed in a few cell types. Discovery of the gasdermin family, which are true cell death executors, has altered our understanding of programmed cell death. In this work, we explored whether members of the gasdermin family might respond to granzymes and induce pyroptosis.

RESULTS: The expression of gasdermin B (GSDMB) but no other gasdermins in human embryonic kidney (HEK) 293T cells induced pyroptotic killing by NK cells, accompanied by an interdomain cleavage of GSDMB. These processes were blocked by inhibiting the perforin–granzyme pathway. In vitro profiling

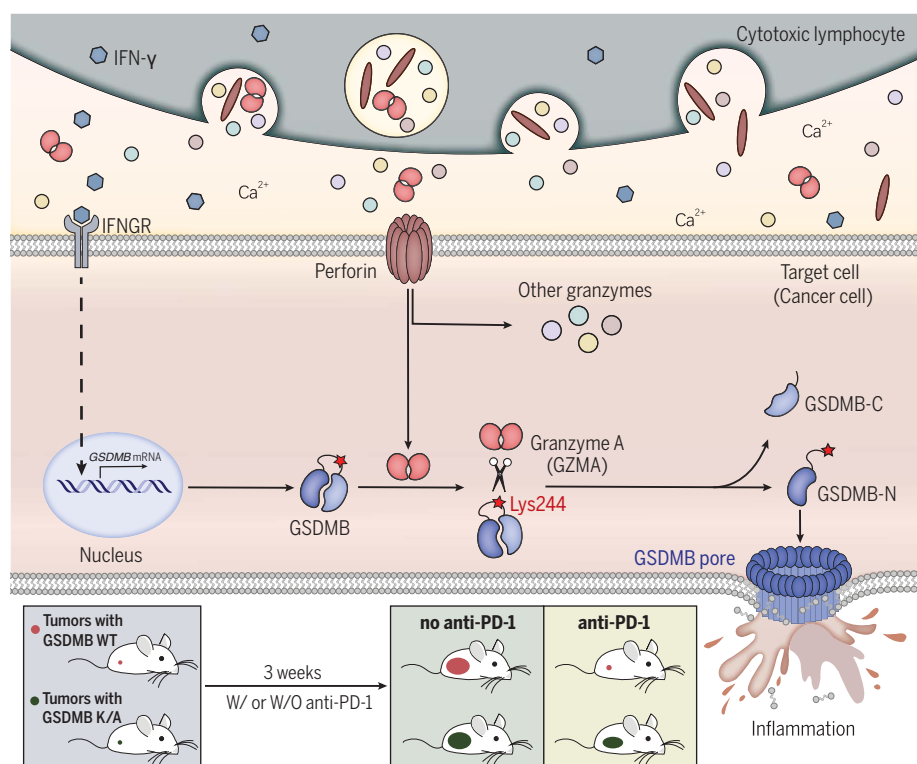
of all five human granzymes identified granzyme A (GZMA), which readily cleaved GSDMB, predominantly at Lys²⁴⁴ within the interdomain linker. This cleavage unmasked the pore-forming activity of GSDMB. GZMA, delivered into GSDMB-reconstituted cells by electroporation or perforin,

ON OUR WEBSITE

Read the full article at <https://dx.doi.org/10.1126/science.aaz7548>

induced extensive pyroptosis with interdomain cleavage of GSDMB. These effects were diminished by a K229A/K244A (KK/AA) mutation of GSDMB [in which lysine (K) was replaced by alanine (A) at positions 229 and 244, respectively]. In cells normally undergoing apoptosis upon GZMA delivery, the additional expression of GZMA-cleavable GSDMB converted apoptosis into pyroptosis. Pyroptotic killing by NK cells was blocked by both the KK/AA mutation and a knockdown of GZMA expression. Among 39 cell lines, three, including the esophageal carcinoma (OE19 cells), expressed GSDMB and underwent pyroptosis upon GZMA delivery. Knockout experiments revealed that pyroptosis in OE19 cells required the interdomain cleavage of GSDMB. Furthermore, GSDMB expression was up-regulated by interferon- γ (IFN- γ). Approximately one-third of GSDMB-negative cell lines showed IFN- γ -induced GSDMB expression. IFN- γ promoted GZMA- or NK cell-induced pyroptosis in several target cells. Primary T cells, including anti-CD19 chimeric-antigen receptor (CAR) T cells and NY-ESO-1-specific T cell receptor (TCR)-engineered T cells (TCR T cells), also induced pyroptosis in GSDMB-reconstituted cells through cleavage of GSDMB by GZMA. Introducing GZMA-cleavable GSDMB into mouse cancer cells promoted tumor clearance in mice. GSDMB was highly expressed in certain tissues, particularly digestive tract epithelia, including the derived tumors. GSDMB appeared to be silenced in gastric and esophageal cancers. The Cancer Genome Atlas database recorded a strong positive correlation between GSDMB expression and patient survival for bladder carcinoma and skin cutaneous melanoma.

CONCLUSION: GZMA from cytotoxic lymphocytes cleaves and activates GSDMB to induce target cell pyroptosis. This immune effector mechanism promotes CTL-mediated tumor clearance in mice. High GSDMB expression in the digestive system suggests the importance of GSDMB-mediated immunity in these tissues and will guide immunotherapy for related cancers. Our findings suggest that substrates such as gasdermins, rather than their upstream proteases, determine the nature of cell death. ■



GZMA from cytotoxic lymphocytes cleaves GSDMB in target cells, predominantly at Lys²⁴⁴ within the interdomain linker. The cleavage allows GSDMB pore-forming domain (GSDMB-N) to perforate plasma membrane and induce pyroptosis. Expression of GSDMB wild type (WT) but not its GZMA-resistant K/A mutant in mouse cancer cells promotes cytotoxic T lymphocyte-mediated tumor clearance when the inhibitory checkpoint is blocked by antibody to programmed cell death 1 (PD-1). IFNGR, IFN- γ receptor.

The list of author affiliations is available in the full article online.

*These authors contributed equally to this work.

†Corresponding author. Email: shaofeng@nibs.ac.cn

Cite this article as Z. Zhou et al., *Science* 368, eaaz7548 (2020). DOI: 10.1126/science.aaz7548

RESEARCH ARTICLE

IMMUNOLOGY

Granzyme A from cytotoxic lymphocytes cleaves GSDMB to trigger pyroptosis in target cells

Zhiwei Zhou^{1,3,*}, Huabin He^{2,3,*}, Kun Wang³, Xuyan Shi³, Yupeng Wang^{1,3}, Ya Su³, Yao Wang⁴, Da Li³, Wang Liu^{1,3}, Yongliang Zhang⁵, Lianjun Shen⁵, Weidong Han⁶, Lin Shen⁶, Jingjin Ding^{3,7}, Feng Shao^{1,3,7,8,†}

Cytotoxic lymphocyte-mediated immunity relies on granzymes. Granzymes are thought to kill target cells by inducing apoptosis, although the underlying mechanisms are not fully understood. Here, we report that natural killer cells and cytotoxic T lymphocytes kill gasdermin B (GSDMB)-positive cells through pyroptosis, a form of proinflammatory cell death executed by the gasdermin family of pore-forming proteins. Killing results from the cleavage of GSDMB by lymphocyte-derived granzyme A (GZMA), which unleashes its pore-forming activity. Interferon- γ (IFN- γ) up-regulates GSDMB expression and promotes pyroptosis. GSDMB is highly expressed in certain tissues, particularly digestive tract epithelia, including derived tumors. Introducing GZMA-cleavable GSDMB into mouse cancer cells promotes tumor clearance in mice. This study establishes gasdermin-mediated pyroptosis as a cytotoxic lymphocyte-killing mechanism, which may enhance antitumor immunity.

Cytotoxic lymphocytes, primarily cytotoxic T lymphocytes (CTLs) and natural killer (NK) cells, are major immune effectors that kill virus-infected or transformed cells or grafted tissues. This cell-mediated cytotoxicity is downstream of either the Fas death receptor or, more commonly, the granule exocytosis pathway. This second pathway relies on perforin to deliver serine proteases called granzymes into target cells (1, 2). Understanding how cytotoxic lymphocytes use these two pathways to kill target cells is a central question in immunology, particularly given the key role of CTLs in cancer immunotherapy. Extensive studies beginning in the 1990s (3–6) have fed the predominant view that both the Fas and the perforin-granzyme pathways induce target cell apoptosis (7, 8). However, these earlier studies were performed at the time when apoptosis was considered to be the only or dominant form of programmed

cell death. Assays to ascertain apoptosis were not particularly accurate and were conducted in vitro in limited types of target cells. There have been occasional reports indicating non-apoptotic cell death induced by cytotoxic lymphocytes. For instance, when targeted by NK cells, K562 and CHP134 cells undergo necrosis (9, 10), and Jurkat cells show a mixed form of cell death (11). Nonapoptotic killing by CTLs has also been noted under certain contexts (12–15). Thus, the nature of lymphocyte cytotoxicity may be context-dependent.

The granzyme family comprises five members in humans—granzyme A (GZMA), granzyme B (GZMB), granzyme H (GZMH), granzyme K (GZMK), and granzyme M (GZMM)—and 11 in mice, each with its own substrate specificity. Granzymes have been studied for more than 30 years (1), with the major focus placed on the two most abundant members (GZMA and GZMB) and their role in killing cancer cells (16). GZMB is believed to induce target cell apoptosis by cleaving caspase-3 or caspase-3 substrates (3, 6, 16). The action of GZMA is less clear. In vitro studies using perforin-loaded GZMA have identified diverse substrates in single-strand DNA damage and mitochondria electron transport complex. This has led investigators to propose a sophisticated caspase-independent apoptosis pathway whose mechanism has not been defined (16, 17). The contribution of the known cleavage activities of GZMA to CTL- and NK cell-mediated killing is yet to be established. Furthermore, *Gzma*^{−/−} mice (18) reveal no apparent defects in apoptosis, and mouse GZMA can induce a nonapoptotic form of cell death with writhing morphology (19). Thus, the actual function and mechanism of action of GZMA remain unclear, likely because

of the lack of sufficiently relevant experimental systems.

Pyroptosis, a form of proinflammatory cell death, is emerging as a critical effector mechanism in innate immunity. Historically, pyroptosis was regarded as caspase-1-mediated monocyte death, accompanied by interleukin-1 β (IL-1 β) and IL-18 secretion (20). Recent work shows that mouse caspase-11 as well as its human homologs caspase-4 and -5 also induce pyroptosis upon direct activation by lipopolysaccharide (LPS) (21, 22). Additionally, gasdermin D (GSDMD) has been shown to be the pyroptosis executor downstream of both caspase-1 and caspase-11/4/5 (23, 24). Consequently, pyroptosis is now redefined as gasdermin-mediated programmed necrosis (25, 26).

GSDMD bears an N-terminal pore-forming domain and a C-terminal inhibitory domain. Interdomain cleavage of GSDMD by caspase-1 or caspase-11/4/5 allows the N-terminal domain to bind membrane phospholipids and oligomerize into a large ~18-nm pore on the plasma membrane (27–30). Supporting the redefinition of pyroptosis, four other gasdermins (GSDMA, GSDMB, GSDMC, and GSDME) share the autoinhibited pore-forming domain but not the caspase-1/11 cleavage site (24, 28, 31). GSDME, the other well-characterized gasdermin, is cleaved and activated by caspase-3 similarly as GSDMD by caspase-1/11 (32, 33). GSDME expression can be appreciably detected in only ~10% of cancer cells (33–37). GSDME can switch caspase-3-mediated cell death from apoptosis to pyroptosis, rendering cancer cells more sensitive to killing agents (33). These findings support the early argument that GSDME may act as a tumor suppressor. Beyond GSDMD and GSDME, the mechanisms of activation are unknown for other members of the gasdermin family.

NK cells induce GSDMB-dependent pyroptosis through interdomain cleavage of GSDMB

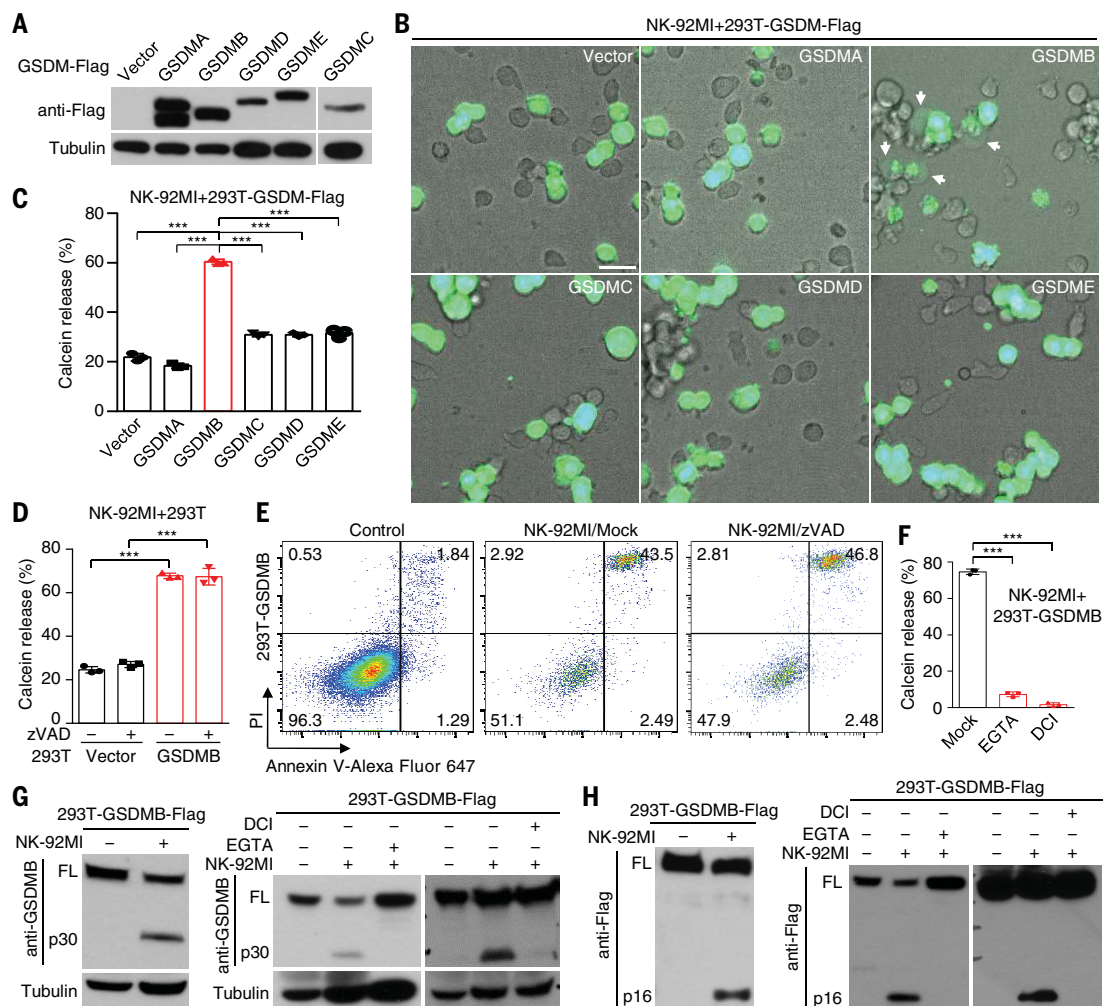
We first asked whether gasdermins might play a role in lymphocyte cytotoxicity. Because most gasdermins have cell type-selective expression and certain gasdermins are silenced in cancers, we reconstituted the expression of GSDMA, GSDMB, GSDMC, GSDMD, and GSDME individually in human embryonic kidney (HEK)-293T cells (Fig. 1A), which express no endogenous gasdermins. When cocultured with NK-92MI cells, an established human NK cell line, 293T cells harboring the empty vector showed little cell death, despite occasional apoptosis, which is consistent with previous reports (Fig. 1B and movie S1) (38). Cells expressing GSDMB, but not any of the other four gasdermins, developed evident pyroptotic morphology (Fig. 1B and movie S2). Annexin V and propidium iodide (PI) staining confirmed that NK-92MI cells induced more extensive lytic death in

¹Research Unit of Pyroptosis and Immunity, Chinese Academy of Medical Sciences and National Institute of Biological Sciences, Beijing, Beijing 102206, China. ²Peking University–Tsinghua University–National Institute of Biological Sciences Joint Graduate Program, School of Life Sciences, Tsinghua University, Beijing 100084, China. ³National Institute of Biological Sciences, Beijing, Beijing 102206, China. ⁴Department of Molecular and Immunology and Department of Bio-therapeutics, Chinese PLA General Hospital, Beijing 100853, China. ⁵Gracell Biotechnologies, Shanghai 200233, China. ⁶Department of Gastrointestinal Oncology, Peking University Cancer Hospital and Institute, Beijing 100142, China. ⁷National Laboratory of Biomacromolecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences (CAS), Beijing 100101, China. ⁸Tsinghua Institute of Multidisciplinary Biomedical Research, Tsinghua University, Beijing 102206, China. *These authors contributed equally to this work. †Corresponding author. Email: shaofeng@nibs.ac.cn

Fig. 1. GSDMB expressed in target cells confers pyroptotic killing by NK cells and undergoes inter-domain cleavage.

(A) Immunoblotting of gasdermin-Flag (GSDM-Flag) expression in 293T cells. **(B)** Representative images of gasdermin-expressing 293T cells after a 6-hour incubation with NK-92MI cells. 293T cells were preloaded with a fluorescent dye calcein AM. Arrowheads indicate pyroptotic cells. Scale bar, 25 μ m. **(C)** Quantification of pyroptosis by measuring calcein release from the

gasdermin-expressing 293T cells after a 6-hour incubation with NK-92MI cells. **(D)** Effect of the pan-caspase inhibitor zVAD on NK-92MI cell-induced calcein release from GSDMB-expressing 293T cells. **(E)** Representative flow cytometry pseudocolor dot-plots of PI- and annexin V-stained GSDMB-expressing 293T cells after incubation with NK-92MI cells in the presence or absence of zVAD. **(F)** Effect of EGTA and a pan-serine protease inhibitor DCI on NK-92MI cell-induced calcein release from GSDMB-expressing 293T cells. **(G and H)** Immunoblotting of GSDMB-Flag cleavage in NK-92MI cell-treated 293T cell and inhibition by EGTA and DCI. The antibodies to GSDMB and Flag recognize the N- and C-terminal cleavage products, respectively. Data are means $\pm$ SD from three replicates [(C), (D), and (F)]. *** P < 0.001; two-tailed unpaired Student's t test. Data are representative of three [(A to D) and (F)] or two [(E), (G), and (H)] independent experiments.



GSDMB-expressing 293T cells than that in cells expressing other gasdermins or the empty vector (fig. S1, A to C).

To specifically examine target cell lysis in the coculture killing assay, we preloaded 293T cells with the calcein AM (acetoxymethyl) dye and performed the classical calcein release assay. Despite low background killing by NK-92MI cells in control 293T cells, this more sensitive assay confirmed that the expression of GSDMB, but not any of the other four gasdermins, induced extensive lysis of 293T cells (Fig. 1, B and C, and fig. S1D). Consistent with the inability of GSDMD and GSDME to sensitize 293T cells to killing by NK-92MI cells, these two gasdermins were not cleaved (fig. S1, E to G). Compared with 293T cells, HeLa and human melanoma A375 cells expressed higher amounts of caspase-3 and were more sensitive to NK-92MI cells owing to caspase-3 cleavage of GSDME. However, additional expression of GSDMB in these two

cell lines still promoted killing by NK-92MI cells (fig. S1, H to K). Both flow cytometry and calcein release assays further revealed that GSDMB-rendered pyroptotic killing by NK cells was completely resistant to the pan-caspase inhibitor zVAD (carbobenzyloxy-valyl-alanyl-aspartyl-[O-methyl]-fluoromethylketone) (Fig. 1, D and E). Thus, NK cells can induce target cells to undergo GSDMB-dependent but caspase-independent pyroptosis.

We next examined whether GSDMB-mediated pyroptosis is downstream of cytotoxic granule-derived perforin and granzyme activity. The Ca^{2+} chelator EGTA, an inhibitor of perforin and NK cell degranulation (39), blocked NK-92MI-induced pyroptosis in GSDMB-expressing 293T cells (Fig. 1F). A similar effect was observed with the pan-granzyme inhibitor DCI (3,4 dichloroisocoumarin) (Fig. 1F) (40). Consistent with the requirement of granzyme activity, immunoblotting by using an antibody that recognizes the N-terminal domain of GSDMB

revealed a ~30-kDa (p30) cleavage fragment (Fig. 1G). The presence of a Flag tag at the carboxyl terminus of GSDMB allowed identification of another ~16-kDa (p16) fragment in NK-92MI-treated 293T cells (Fig. 1H). The appearance of these two cleavage fragments, whose combined molecular weight equaled that of full-length GSDMB (~47 kDa), was diminished by the addition of EGTA or DCI to the coculture killing assay (Fig. 1, G and H). Thus, GSDMB-dependent pyroptosis is due to intrinsic killing activity of NK cells and requires the specific cleavage of GSDMB.

Site-specific cleavage of GSDMB by GZMA releases its pore-forming activity

The above results strongly indicate that a granzyme mediates the interdomain cleavage of GSDMB and induces target cell pyroptosis. To test this hypothesis, we purified all five human granzymes from 293F cells whose activities were confirmed through in vitro cleavage

of their known substrates such as SET and DFF45 (fig. S2, A and B) (41–45). GZMA, but not the other four granzymes, readily cleaved purified GSDMB, with a cleavage pattern akin to that observed in NK cell-killing assays (Fig. 2A). Purified GSDMD resisted cleavage by all five granzymes (fig. S2C). As expected, a mutation of the catalytic Ser²¹² residue in GZMA abolished its ability to cleave GSDMB (fig. S2D). At high doses of GZMA, near-complete cleavage was observed, resulting in two major products (p30 and p16) that remained stable

(fig. S2E). Under these conditions, we observed another pair of minor cleavage products (p28 and p18), indicating a second but much less preferred cleavage site. Edman sequencing of the N termini of the p16 and p18 species identified the cleavage sites to be Lys²⁴⁴ and Lys²²⁹ (Fig. 2B and fig. S2E). This was consistent with the fact that GZMA is a tryptic protease that cleaves after a lysine or arginine (17). Lys²⁴⁴ is shared by all four splicing isoforms of GSDMB, but Lys²²⁹ is only present in isoforms 3 and 4 (the longest isoform 3 was used in this study)

(fig. S2F). The four isoforms showed a nearly identical pattern of cleavage by GZMA (fig. S2G). This supports the notion that Lys²⁴⁴ is the dominant GZMA cleavage site. A mutation of Lys²²⁹ to alanine (K229A) did not block the generation of the p28 and p18 fragments, likely because of alternative cleavage at two adjacent lysine residues (Lys²²⁷ and Lys²²⁵) (fig. S2H). Although GSDMB K229A was still efficiently cleaved at the Lys²⁴⁴ site, the K244A or the K229A/K244A double mutation [in which lysine (K) was replaced by alanine (A) at positions 229 and 244, respectively] drastically inhibited the cleavage by GZMA (fig. S2H). The minor cleavage products only appeared in vitro with excessive GZMA but not during NK cell killing. Thus, Lys²⁴⁴ in GSDMB is the major and physiological cleavage site by GZMA.

Interdomain cleavage of GSDMD and GSDME by caspases unleashes the pore-forming activity of their N-terminal domains. GSDMB has a shorter interdomain linker (residues 221 to 252) covering the GZMA cleavage sites (fig. S2F). Like the situation observed with GSDMD and GSDME, a purified noncovalent complex of GZMA-cleavage products GSDMB-N (residues 1 to 244) and GSDMB-C (residues 245 to 416) [dubbed GSDMB-(N+C)] (fig. S2I) caused severe leakage of phosphatidylinositol-4,5-bisphosphate [PI(4,5)P₂]-containing liposomes (fig. S2J). By contrast, uncleaved full-length GSDMB had no such effect. GSDMB-(N+C) but not intact GSDMB protein could trigger extensive pyroptosis in 293T cells. This effect was only observed when GSDMB-(N+C) protein was delivered into 293T cell cytosol but not added to the extracellular space (Fig. 2C and fig. S2K). Thus, the cleavage of GSDMB by GZMA at Lys²⁴⁴ is sufficient to unmask its pore-forming domain.

GZMA cleavage of GSDMB is required for pyroptotic killing by NK cells

We next examined the function of GSDMB cleavage by GZMA. Electroporation of purified GZMA into GSDMB-reconstituted 293T induced extensive cell death with pyroptotic morphologies, whereas control cells harboring an empty vector showed no such responses (fig. S3, A and B). Pyroptotic killing of 293T cells was not observed with the protease-deficient GZMA S212A mutant (fig. S3, A and B). Immunoblotting identified the p30 cleavage product of GSDMB, which was absent in cells electroporated with GZMA S212 (fig. S3C). Consistent with the absence of the p28 fragment, the K229A mutation in GSDMB did not affect the cleavage induced by electroporation of GZMA (fig. S3D). By contrast, this cleavage was diminished in cells expressing GSDMB K244A or the K229A/K244A double mutant (fig. S3D). Accordingly, although GSDMB K229A could still mediate GZMA-induced pyroptosis, cells expressing GSDMB K244A were largely

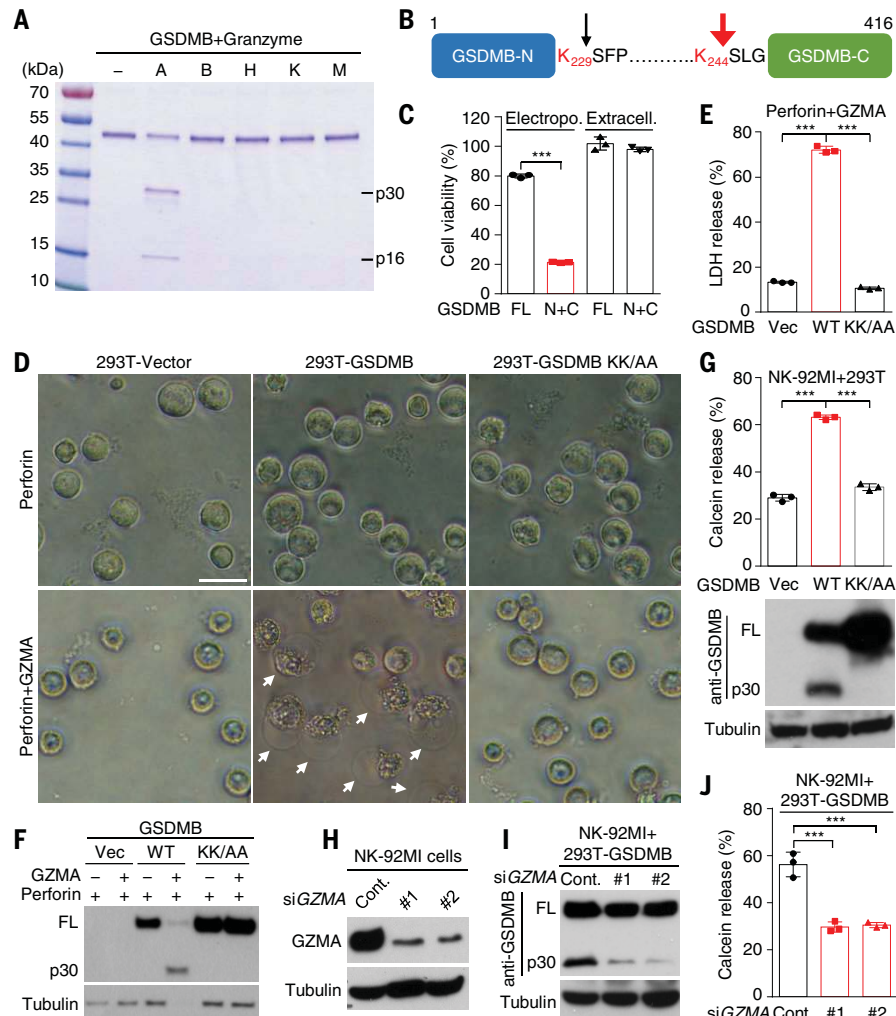


Fig. 2. GZMA cleaves GSDMB to induce pyroptosis in GSDMB-reconstituted target cells. (A) In vitro cleavage assay of GSDMB by five human granzymes. Shown is a representative Coomassie Blue-stained gel. (B) Diagram of GSDMB structure and the GZMA cleavage sites (Lys²⁴⁴ and Lys²²⁹). (C) Equal amounts of full-length (FL) GSDMB or GSDMB-(N+C) proteins were electroporated into or added directly to 293T cells. ATP-based cell viability is shown. (D to F) GZMA was delivered into 293T cells expressing GSDMB WT or K229A/K244A (KK/AA) mutant by perforin. (D) Representative cell images (arrowheads indicate pyroptotic cells). Scale bar, 25 μ m. (E) LDH release-based cell death. (F) Immunoblotting of GSDMB cleavage. (G) 293T cells expressing GSDMB WT or KK/AA double mutant were preloaded with calcein AM and then cocultured with NK-92MI cells. Quantification of calcein AM release and immunoblotting of GSDMB cleavage are shown. (H to J) Effect of siRNA knockdown of GZMA on pyroptotic killing of GSDMB-expressing 293T cells by NK-92MI cells. (H) Anti-GZMA immunoblot in NK-92MI cells. (I) Immunoblotting of GSDMB cleavage. (J) Quantification of calcein AM release. Data are means $\pm$ SD from three replicates [(C), (E), (G), and (J)]. *** P < 0.001; two-tailed unpaired Student's t test. Data are representative of three [(A) and (C to F)] or two [(G) to (J)] independent experiments.

resistant to pyroptosis. Furthermore, complete resistance was observed with 293T cells expressing GSDMB K229A/K244A (fig. S3, E and F). For simplicity, the double mutant was used in subsequent analyses. In a human colon cancer cell line (SW480) that naturally expressed GSDMD but no other gasdermins (fig. S4A), electroporation of GZMA triggered extensive apoptotic death (fig. S4, B and C). Although the physiological relevance of this cell death requires further investigation, GZMA-induced apoptosis was converted to pyroptosis by expressing GSDMB wild type (WT) but not its K229A/K244A double mutant in SW480 cells (fig. S4, B and C). In a correlation with the cell death response, the p30 fragment of GSDMB was detected in cells expressing GSDMB WT but not the K229A/K244A double mutant (fig. S4D). Thus, the cleavage of GSDMB at Lys²⁴⁴ by GZMA is necessary and sufficient to trigger pyroptosis.

We also examined GZMA cleavage of GSDMB under physiological settings. Similar to electro-

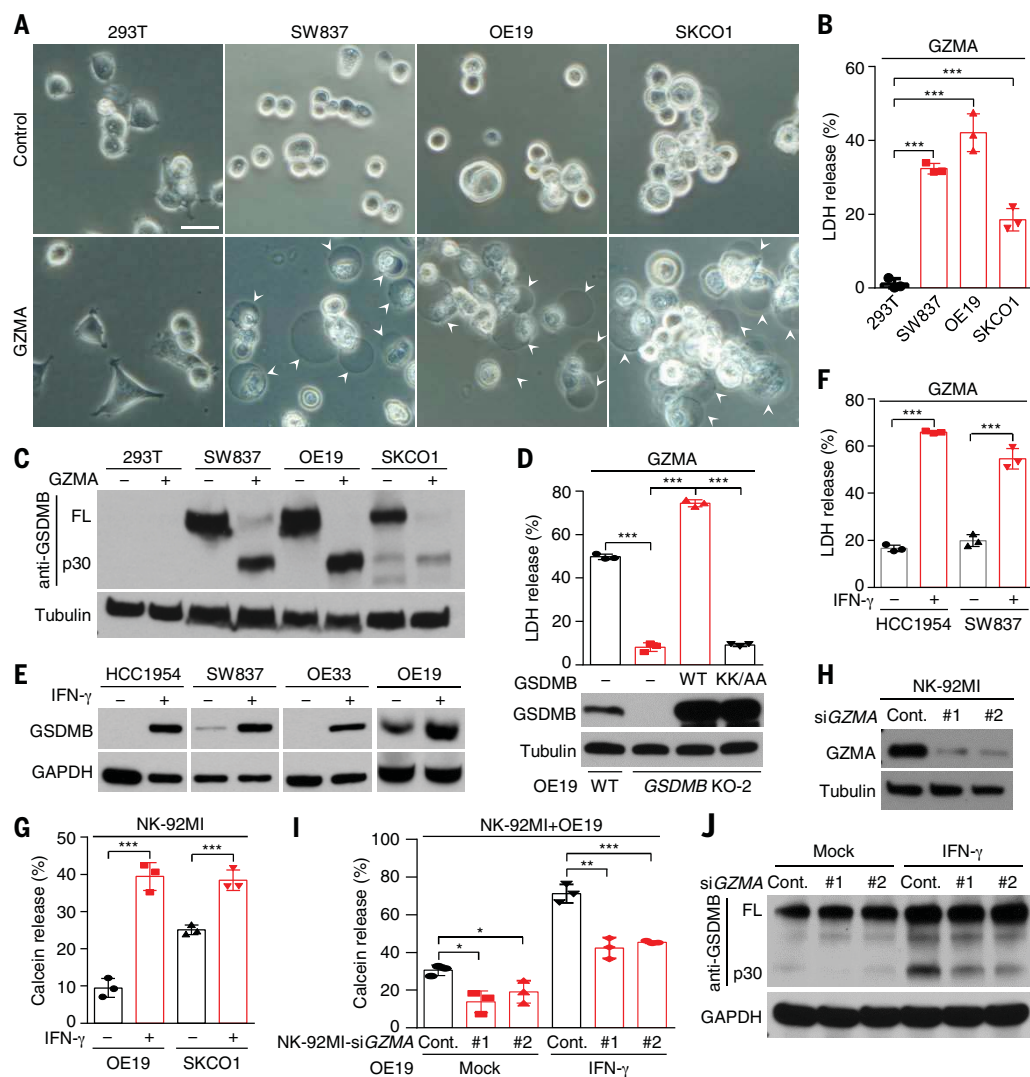
poration of GZMA, cytosolic delivery of GZMA by purified perforin caused extensive pyroptosis in GSDMB-expressing 293T cells. This did not occur with perforin alone and was diminished in cells expressing the GSDMB K229A/K244A double mutant (Fig. 2, D and E). Accordingly, GSDMB WT but not K229A/K244A was cleaved into the p30 fragment in response to perforin-delivered GZMA (Fig. 2F). In the NK cell-killing assay, blocking GZMA cleavage by the K229A/K244A double mutation in GSDMB also reduced pyroptosis in 293T cells (Fig. 2G). Accordingly, small interfering RNA (siRNA) knockdown of *GZMA* in NK-92MI cells inhibited the cleavage of GSDMB as well as pyroptosis in GSDMB-expressing 293T cells (Fig. 2, H to J). YTS is another NK cell line that expresses little GZMA endogenously (fig. S4E). When GZMA WT but not S212A mutant was reconstituted into YTS cells, the cells gained the ability to induce pyroptosis in GSDMB-expressing 293T cells, accompanied by interdomain cleavage of GSDMB (fig. S4, E to G). Thus, cleavage of

GSDMB by GZMA causes pyroptotic killing of target cells by NK cells.

Endogenous GSDMB is sufficient to mediate GZMA-induced pyroptosis

To further probe the physiological role of GZMA cleavage of GSDMB, we sought to examine cells that express GSDMB endogenously. Among a panel of 39 different human cancer cell lines, three of them—OE19 (esophageal carcinoma), SW837 (rectum adenocarcinoma), and SKCO1 (colorectal adenocarcinoma)—showed apparent GSDMB expression, whereas the remaining 36 cell lines expressed little or no GSDMB (fig. S5A). When active GZMA was electroporated into the three GSDMB⁺ cells, around 40% of SW837 and OE19 cells and 20% of SKCO1 cells underwent lytic death, featuring pyroptotic morphologies (Fig. 3, A and B). Furthermore, robust interdomain cleavage of endogenous GSDMB—evident by the appearance of the p30 fragment—appeared concurrently (Fig. 3C).

Fig. 3. Endogenous GSDMB mediates GZMA-dependent killing, and IFN- γ primes GSDMB expression to promote this killing. (A to D) GZMA cleavage of GSDMB-induced pyroptosis in GSDMB⁺ cancer cells. GZMA protein was electroporated into SW837, OE19, or SKCO1 cells. (A) Representative cell images. Arrowheads indicate pyroptotic cells. Scale bar, 25 μ m. [(B) and (D)] LDH release-based cell death. [(C) and (E)] Anti-GSDMB immunoblotting. GSDMB WT or K229A/K244A (KK/AA) mutant was expressed in GSDMB^{-/-} OE19 cells (KO-2) in (D). (E to G) Effect of IFN- γ on GSDMB expression and GZMA or NK cell-induced pyroptosis. HCC1954, SW837, OE33, and OE19 cells were treated with IFN- γ for 48 hours. Immunoblotting of GSDMB expression is shown in (E). GZMA protein was electroporated into IFN- γ -treated or -untreated HCC1954 and SW837 cells, and cell death was measured by LDH release in (F); IFN- γ -treated or -untreated OE19 and SKCO1 cells were cocultured with NK-92MI cells, and cell death was measured by calcein release in (G). (H to J) Effect of siRNA knockdown of GZMA on pyroptotic killing of OE19 cells by NK-92MI cells. (H) Anti-GZMA immunoblot in NK-92MI cells. (I) Quantification of calcein release. (J) Immunoblotting of GSDMB cleavage. Data are means $\pm$ SD from three replicates [(B), (D), (F), (G), and (I)]. *** P < 0.001, ** P < 0.01, * P < 0.05; two-tailed unpaired Student's t test. Data are representative of three [(A) to (F) and (H) to (J)] or two (G) independent experiments.



The knockdown of *GSDMB* expression in OE19 cells by any of the three independent *GSDMB*-specific siRNAs effectively blocked GZMA electroporation-induced pyroptosis (fig. S6A). *GSDMB*-knockout (KO) OE19 cells generated by using the CRISPR-Cas9 method also resisted pyroptotic killing by GZMA (fig. S6B). The reexpression of *GSDMB* WT but not the K229A/K244A double mutant in the *GSDMB*^{-/-} OE19 cells restored GZMA electroporation-induced pyroptosis (Fig. 3D). When GZMA was delivered into OE19 cells by perforin, endogenous *GSDMB* underwent interdomain cleavage with the appearance of the p30 fragment (fig. S6C). Perforin-delivered GZMA also induced pyroptosis in OE19 cells, which was diminished in all three *GSDMB* KO clones (fig. S6, D and E). Thus, *GSDMB*⁺ cells undergo pyroptosis in response to interdomain cleavage by GZMA.

Interferon- γ up-regulates *GSDMB* expression and promotes GZMA-induced pyroptosis

Given the absence of detectable *GSDMB* protein in many cell lines (fig. S5A), we investigated whether *GSDMB* expression could be transcriptionally induced. SW837 cells were individually treated with a panel of cytokine receptor agonists or kinase inhibitors (Fig. 3E and fig. S5B). Interferon- α (IFN- α), IFN- β , IFN- γ , and to a lesser extent tumor necrosis factor- α (TNF- α) increased the expression of endogenous *GSDMB*. Although the action of TNF- α was specific to certain cells, IFN- γ exhibited broad effects: 11 of the 36 *GSDMB*⁻ cells in our initial profiling showed evidently increased *GSDMB* expression (fig. S5C). IFN- γ also up-regulated *GSDMB* expression in several other cell lines, including two esophageal carcinoma cell lines (OE19 and OE33) and a breast cancer cell line (HCC1954). HCC1954 and OE33 cells normally expressed little to no *GSDMB* (Fig. 3E).

Both IFN- γ and TNF- α are released by activated cytotoxic lymphocytes, and the former critically promotes or drives cell-mediated immunity. Up-regulation of *GSDMB* by IFNs and TNF- α underscores its role in lymphocyte cytotoxicity. IFN- γ priming of HCC1954 and SW837 cells markedly promoted pyroptosis induced by the cytosolic delivery of GZMA (Fig. 3F and fig. S6F). In the NK-92MI cell-killing assay, pretreating SKCO1 or OE19 cells with IFN- γ also caused more extensive pyroptotic death (Fig. 3G). In both assays (Fig. 3, F and G), possible cell death caused by IFN- γ alone (fig. S6G) was excluded from the analyses by extensive washing performed between IFN- γ priming and GZMA-mediated killing. Both basal and IFN- γ -primed pyroptosis in SKCO1 or OE19 cells were inhibited by siRNA knockdown of *GZMA* in the NK-92MI cells (Fig. 3, H and I). Knockdown of *GZMA* also reduced the interdomain cleavage of IFN- γ -primed endogenous *GSDMB* (Fig. 3, H and J). Thus,

GSDMB expression can be induced by IFN- γ , which promotes the GZMA-mediated pyroptotic killing of target cells.

GZMA cleavage of *GSDMB* mediates pyroptotic killing by CTLs

Similar to NK cells, CTLs use the perforin-granzyme pathway to confer cytotoxicity on target cells and also express high amounts of GZMA (2). To determine whether GZMA cleavage of *GSDMB*-induced pyroptosis contributes to CTL cytotoxicity, we first assayed the killing of CD19-expressing 293T cells by human anti-CD19 chimeric-antigen receptor (CAR) T cells. In a calcein release assay, the CAR T cells caused severe leakage of calcein fluorescence from the 293T cells only when they expressed *GSDMB* (Fig. 4, A and B). Cells that released calcein exhibited evident pyroptotic morphologies. CAR T cell-induced pyroptosis was not detected in 293T cells expressing the K229A/K244A double mutant of *GSDMB* (Fig. 4, A and B). Accordingly, interdomain cleavage in response to incubation with the CAR T cells occurred only with WT but not the K229A/K244A double mutant of *GSDMB* (Fig. 4B). Conversely, siRNA knockdown of *GZMA* in CAR T cells inhibited calcein release from *GSDMB*-expressing 293T cells (Fig. 4, C and D).

We also assayed primary human T cells engineered to express a specific T cell receptor (TCR) that recognizes the immunogenic tumor antigen NY-ESO-1. These TCR-engineered T cells (TCR T cells) induced evident pyroptosis in NY-ESO-1⁺ A375 cells when *GSDMB* was co-expressed (Fig. 4, E and F). These TCR T cells also stimulated the interdomain cleavage of *GSDMB* but did not activate the caspase-3-GSDME pathway because of the low amounts of GZMA in the particularly primary T cells (Fig. 4, G and H). The K229A/K244A mutation in *GSDMB* not only blocked its interdomain cleavage but also rendered A375 cells resistant to pyroptotic killing by TCR T cells (Fig. 4, E to G).

Unlike other human gasdermins, *GSDMB* has no orthologs in mice. However, mouse GZMA (mGZMA), sharing ~89% sequence similarity with human GZMA, could also cleave human *GSDMB* into the p30 and p16 fragments, albeit with less activity than that of human GZMA (fig. S7A). The cleavage by mGZMA was also diminished by the K229A/K244A double mutation in *GSDMB*. Consistently, the cytosolic delivery of mGZMA induced pyroptosis in 293T cells that express *GSDMB* WT but not the K229A/K244A double mutant (fig. S7B). This correlated with the interdomain cleavage of the expressed *GSDMB* (fig. S7C). We then isolated splenocytes from OT-1 transgenic mice and stimulated them with the SIINFEKL (OVA²⁵⁷⁻²⁶⁴) peptide to generate CTLs. In a co-culture killing assay, these splenic CTLs triggered evident *GSDMB*-dependent pyroptosis

in mouse colon adenocarcinoma MC38 cells that had been pulsed with the OVA²⁵⁷⁻²⁶⁴ peptide (Fig. 4I). This pyroptosis response was not observed in MC38 cells expressing the K229A/K244A double mutant of *GSDMB* (Fig. 4I). Thus, CTLs can also use GZMA to cleave *GSDMB* in target cells and therefore kill them through pyroptosis.

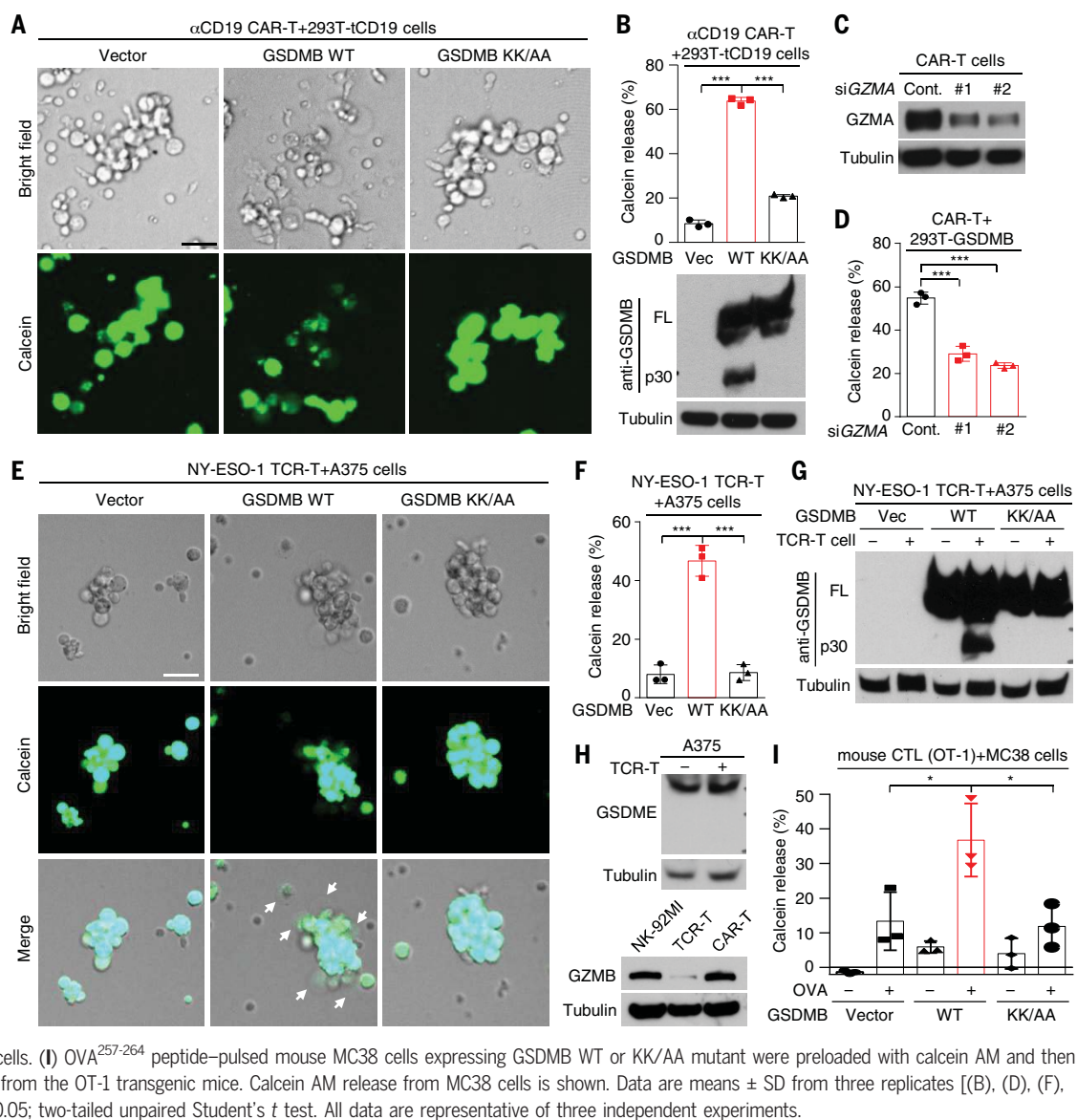
GZMA cleavage of *GSDMB* promotes tumor clearance in mice

Granzyme-mediated cell death is a critical mechanism for cytotoxic lymphocytes to eliminate malignant cells in antitumor immunity. We therefore hypothesized that GZMA-induced and *GSDMB*-dependent pyroptosis may have an important function in CTL-mediated tumor clearance, particularly given the proinflammatory nature of pyroptosis. To test this hypothesis, mouse colon carcinoma CT26 cells reconstituted with human *GSDMB* WT, *GSDMB* K229A/K244A, or an empty vector were engrafted subcutaneously into BALB/c mice (Fig. 5A). Growth of the CT26 tumors was then followed continuously for 3 weeks. However, no statistically significant differences in tumor growth were observed among the three groups of mice; *GSDMB*⁺ and *GSDMB*⁻ tumor grafts showed comparable increases in their volume and weight (Fig. 5, B to E). CT26 tumors are known to respond well to blockade of the programmed cell death 1 (PD-1)-programmed-death ligand 1 (PD-L1) pathway (46–49). The PD-1-PD-L1 interaction inhibits T cell recognition of target cells. Thus, the presence of the PD-1-PD-L1 checkpoint should prevent downstream granzyme functioning in the intact CT26 model. We therefore intraperitoneally injected antibody to PD-1 at days 8, 11, and 14 into the tumor-bearing mice (Fig. 5A). Under this condition, a partial inhibition of tumor growth was observed with the control CT26 tumor grafts (Fig. 5, B to E, and fig. S8A). Expression of *GSDMB* in CT26 cells nearly completely suppressed the tumor growth, and at day 22, the tumor burden was almost negligible (Fig. 5, B to E, and fig. S8A). Expression of the GZMA-resistant K229A/K244A double mutant of *GSDMB* in CT26 cells showed no such effect. Moreover, tumors that express *GSDMB* WT but not the K229A/K244A double mutant contained increased numbers of PI⁺ necrotic cells (fig. S8B).

We also examined the antitumor immune function of *GSDMB* in a more aggressive B16-F10 tumor model. For this, B16-F10 cells expressing an empty vector, *GSDMB* WT, or *GSDMB* K229A/K244A were engrafted subcutaneously into C57BL/6 mice, which was also followed by treatment with antibody to PD-1. Compared with the vector control, expression of *GSDMB* resulted in much slower increases in the tumor volume and the tumor weight, which were not observed with *GSDMB* K229A/K244A double mutant (Fig. 5, F and G,

Fig. 4. GZMA cleavage of GSDMB mediates pyroptotic killing by CTLs. (A to D) CD19-reconstituted 293T cells expressing GSDMB WT or K229A/K244A (KK/AA) double mutant were pre-loaded with calcein AM and then cocultured with anti-CD19 CAR T cells. (A) Representative 293T cell images. Scale bar, 25 μ m. (B) Calcein AM release and immunoblotting of GSDMB cleavage in 293T cells. [(C) and (D)] The CAR T cells were transfected with GZMA-specific siRNA before the killing, and anti-GZMA immunoblot and calcein release assay of 293T cell pyroptosis are shown, respectively.

(E to H) NY-ESO-1 antigen⁺ A375 cells expressing GSDMB WT or KK/AA mutant were preloaded with calcein AM and then cocultured with NY-ESO-1-specific human TCR T cells. [(E) and (F)] Representative A375 cell images and calcein AM release assay of cell pyroptosis, respectively. Scale bar, 25 μ m. [(G) and (H)] Anti-GSDMB and anti-GSDME immunoblotting in A375 cells and anti-GZMB blotting in TCR T and CAR T cells. (I) OVA²⁵⁷⁻²⁶⁴ peptide-pulsed mouse MC38 cells expressing GSDMB WT or KK/AA mutant were preloaded with calcein AM and then cocultured with CTLs isolated from the OT-1 transgenic mice. Calcein AM release from MC38 cells is shown. Data are means $\pm$ SD from three replicates [(B), (D), (F), and (I)]. *** P < 0.001, * P < 0.05; two-tailed unpaired Student's t test. All data are representative of three independent experiments.



and fig. S8A). Theoretically, the antigenicity difference between GSDMB WT and GSDMB K229A/K244A may account for the contrasting tumor clearance effect, but the probability is extremely low. Thus, pyroptosis mediated by the GZMA–GSDMB pathway could potentially serve as an important effector mechanism in antitumor immunity.

Insights into cancer and cancer immunotherapy

Our initial untargeted profiling only identified a few cancer cell lines that were positive for GSDMB expression (fig. S5A). Among a panel of 27 diverse human tissue samples, we recorded high amounts of GSDMB expression in the esophageal epithelium, the tongue, the gastric mucosa, the duodenal mucosa, the jejunal mucosa, the ileal mucosa, the appendix, the colonic mucosa, the rectal mucosa, the trachea, and the bladder (fig. S9A).

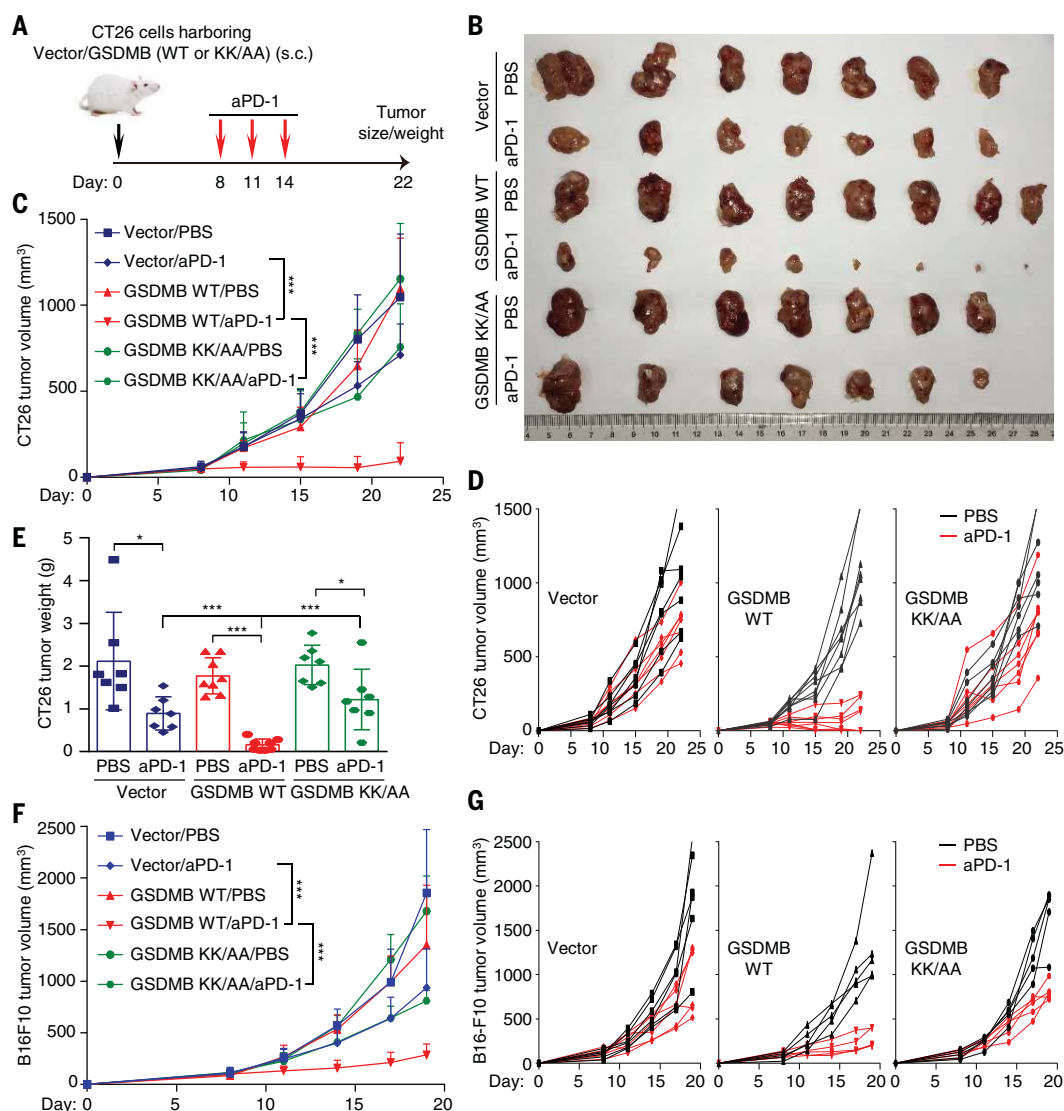
This suggests that GSDMB may play important pathophysiological functions in these tissues. Upon realizing this tissue tropism, we were able to identify 15 more cancer cell lines with high GSDMB expression (fig. S9B). Among these, six were derived from esophageal and gastric tissues—echoing a previous report (50)—eight were from intestinal tissues, and one was ovarian in origin.

We further profiled a collection of diverse clinical tumor samples and found that GSDMB was prevalently expressed in colon, rectal, pancreatic, and cervical cancers but showed little or no expression in breast, lung, and liver cancers (fig. S10, A to D). Amounts of GSDMB expression in the four positive tumors did not significantly differ from those observed in corresponding tumor-adjacent normal tissues. GSDMB amounts in certain colonic, rectal, and cervical tumor tissues were as high as GSDMB exogenously expressed in CT26 and

B16-F10 cells (fig. S10A). We also detected GSDMB expression in 34 out of 75 gastric and 44 out of 80 esophageal tumor tissues (fig. S10, B and C). By contrast, the vast majority of normal tissues corresponding to the 75 gastric and 80 esophageal tumor tissues were GSDMB⁺, suggesting the possible silencing of GSDMB in these two cancer types.

Given that many cancers were positive for GSDMB to varying degrees, we mined the public cancer gene expression and prognosis database using the newly developed GEPIA2 web-server that analyzes the data of 9736 tumors from The Cancer Genome Atlas (TCGA) database (51). Among 33 different tumors analyzed (fig. S11A), GSDMB mRNA expression and its corresponding prognosis value showed a statistically highly significant correlation (P < 0.001) for bladder carcinoma (BLCA), skin cutaneous melanoma (SKCM), and kidney renal clear carcinoma (KIRC) (fig. S11B). In BLCA and

Fig. 5. GZMA cleavage of GSDMB promotes antitumor immunity. (A to E) BALB/c mice were implanted subcutaneously (s.c.) with mouse colon cancer CT26 cells harboring an empty vector or expressing GSDMB WT or the K229A/K244A (KK/AA) double mutant ($n = 7$ or 8 mice per group). (A) The experimental design. [(B) and (E)] Photographs and weights of the CT26 tumors on day 22, respectively. (F and G) B16-F10 cells harboring an empty vector or expressing GSDMB WT or KK/AA mutant ($n = 5$ mice per group) were subcutaneously implanted into C57BL/6 mice. The mice were treated or not treated with antibody to PD-1 (aPD-1) same as that in the CT26 model. [(C) and (F)] Average tumor volumes of each group of mice. [(D and G)] The tumor volume of individual mouse at indicated time points. Data are means $\pm$ SD [(C), (E), and (F)]. * $P < 0.05$, *** $P < 0.001$; two-tailed unpaired Student's t test in (E), two-way analysis of variance (ANOVA) in (C) and (F). Data are representative of three [(B) to (E)] or two [(F) and (G)] independent experiments.



SKCM patients, high *GSDMB* expression was associated with an overall survival benefit, whereas KIRC patients showed a negative correlation. We also examined other gasdermins using the same parameters and observed a positive correlation with low-grade glioma (LGG) and SKCM for *GSDMC* and *GSDMD*, respectively (fig. S11, C and D). Considering the crucial function of *GSDMB* in cytotoxic lymphocyte-mediated tumor clearance, these analyses may prove instructive for future cancer immunotherapy approaches.

Discussion

We have demonstrated that GZMA from cytotoxic lymphocytes cleaves and activates *GSDMB* to induce target cell pyroptosis. The proinflammatory nature of pyroptosis, which distinguishes it from apoptosis, suggests that this type of lymphocyte killing is immunogenic and augments immunity. This ensures a sufficiently robust immune response. The presence of *GSDMB* in various healthy tissues, particu-

larly those composing the digestive system, indicates that the GZMA–*GSDMB* pyroptotic axis may also play important roles in immunity to microbial infections in addition to cancers shown here. GZMA exhibits low cytotoxicity in vitro but considerable proinflammatory properties in vivo (52, 53). This seeming contradiction may be well explained by our finding that GZMA-mediated pyroptotic killing requires *GSDMB*, but that *GSDMB* is not expressed in cells previously used to study GZMA. Moreover, the function of *GSDMB* in cytotoxic lymphocyte-mediated killing is supported by its transcriptional priming by IFN- γ . This sheds new light on our understanding of the multifarious roles played by IFN- γ in immunity, including against tumors.

Last, our finding that *GSDMB* acts downstream of a granzyme reinforces our view that proteolytic enzymes such as caspases as well as granzymes are not cell death executor proteins per se. Rather, these enzymes cleave various downstream substrates, including gasdermins,

which in turn can both initiate and implement particular cell death programs (24, 25, 33). Various target cells, which express a diversity of death executor molecules, may respond differently to the same initiating proteolytic activity, including that of granzymes. This further suggests that different target cells, even when they are similarly recognized by the same type of cytotoxic lymphocyte, may undergo dissimilar types of cell death. Thus, future studies of lymphocyte-mediated immunity, particularly in antitumor immunity, should be more focused on death execution-related events inside target cells rather than surface molecules (such as PD-1).

Methods summary

To explore the possible role played by gasdermins in lymphocyte cytotoxicity, we generated HEK 293T cells that stably expressed each of the five human gasdermins with a C-terminal Flag tag. They were then subjected to a 6-hour incubation with NK-92MI cells. HeLa, A375, and SW480

cells were also used to analyze pyroptosis functions of exogenous GSDMB and its cleavage by GZMA. SW837, OE19, and SKCO1 cells were subjected to GZMA or lymphocyte killing to examine endogenous GSDMB function. HCC1954, SW837, OE19, and SKCO1 cells were used to demonstrate the effect of IFN- γ priming of GSDMB expression and the increased pyroptotic killing by GZMA or cytotoxic lymphocytes. Primary CTLs—including anti-CD19 CAR T cells, NY-ESO-1-specific TCR T cells, and OT-I mice-derived CTLs—were prepared to assess pyroptotic killing of target cells. To evaluate GZMA cytotoxicity, purified GZMA was delivered into target cells either by means of electroporation or perforin. To measure killing as well as to investigate the type of cell death (such as pyroptosis or apoptosis), morphological examination with microscopy, calcein release from target cells, adenosine triphosphate (ATP)-based cell viability and lactate dehydrogenase (LDH) release assays, as well as flow cytometry of annexin V- and PI-stained cells were used where appropriate. siRNA knockdown or CRISPR-Cas9 KO of *GSDMB* in OE19 cells and siRNA knockdown *GZMA* in NK-92MI or CAR T cells were performed to probe the genetic role of GZMA cleavage of GSDMB.

For *in vitro* biochemical experiments, granzymes were expressed and purified from 293F cells. Recombinant gasdermins purified from *E. coli* were incubated with the granzyme to assess possible cleavage, and the cleavage products were visualized on SDS-polyacrylamide gel electrophoresis gels. Throughout the study, a rabbit monoclonal antibody to GSDMB (EPR20841) that recognizes an epitope in the pore-forming domain, developed in collaboration with Abcam, was used for the immunoblot of GSDMB expression in cell lines, immunohistochemistry of GSDMB expression in normal or tumor tissues, as well as immunoblotting of GZMA-catalyzed cleavage during cell killing. To examine GSDMB's antitumor function, GSDMB WT, the K229A/K244A double mutant, or an empty vector were reconstituted into mouse CT26 or B16-F10 cancer cells. The cells were engrafted subcutaneously into BALB/c or C57BL/6 mice. Antibody to PD-1 was then intraperitoneally injected at days 8, 11, and 14 into the tumor-bearing mice. Tumor growth was followed for 3 weeks, and both the tumor volume and weight were measured and subjected to statistical analysis. To evaluate the correlation of gasdermin expression with cancer patient survival, the TCGA database was mined on the GEPIA2 server by using recommended parameters.

REFERENCES AND NOTES

- C. L. Ewen, K. P. Kane, R. C. Bleackley, A quarter century of granzymes. *Cell Death Differ.* **19**, 28–35 (2012). doi: [10.1038/cdd.2011.153](#); pmid: [22052191](#)
- I. Voskoboinik, J. C. Whistock, J. A. Trapani, Perforin and granzymes: Function, dysfunction and human pathology. *Nat. Rev. Immunol.* **15**, 388–400 (2015). doi: [10.1038/nri3839](#); pmid: [25998963](#)
- J. W. Heusel, R. L. Wesselschmidt, S. Shresta, J. H. Russell, T. J. Ley, Cytotoxic lymphocytes require granzyme B for the rapid induction of DNA fragmentation and apoptosis in allogeneic target cells. *Cell* **76**, 977–987 (1994). doi: [10.1016/0092-8674\(94\)90376-X](#); pmid: [8137431](#)
- D. Kagi et al., Fas and perforin pathways as major mechanisms of T cell-mediated cytotoxicity. *Science* **265**, 528–530 (1994). doi: [10.1126/science.7518614](#); pmid: [7518614](#)
- B. Lowin, M. Hahne, C. Mattmann, J. Tschopp, Cytolytic T-cell cytotoxicity is mediated through perforin and Fas lytic pathways. *Nature* **370**, 650–652 (1994). doi: [10.1038/370650a0](#); pmid: [7520535](#)
- A. J. Darmon, D. W. Nicholson, R. C. Bleackley, Activation of the apoptotic protease CPP32 by cytotoxic T-cell-derived granzyme B. *Nature* **377**, 446–448 (1995). doi: [10.1038/377446a0](#); pmid: [7566124](#)
- J. H. Russell, T. J. Ley, Lymphocyte-mediated cytotoxicity. *Annu. Rev. Immunol.* **20**, 323–370 (2002). doi: [10.1146/annurev.immunol.20.100201.131730](#); pmid: [11861606](#)
- J. A. Trapani, M. J. Smyth, Functional significance of the perforin/granzyme cell death pathway. *Nat. Rev. Immunol.* **2**, 735–747 (2002). doi: [10.1038/nri911](#); pmid: [12360212](#)
- Y. Oshimi, K. Oshimi, S. Miyazaki, Necrosis and apoptosis associated with distinct Ca²⁺ response patterns in target cells attacked by human natural killer cells. *J. Physiol.* **495**, 319–329 (1996). doi: [10.1113/jphysiol.1996.sp021596](#); pmid: [8887746](#)
- S. S. Somanchi, K. J. McCulley, A. Somanchi, L. L. Chan, D. A. Lee, A novel method for assessment of natural killer cell cytotoxicity using image cytometry. *PLoS ONE* **10**, e0141074 (2015). doi: [10.1371/journal.pone.0141074](#); pmid: [26492577](#)
- C. S. Backes et al., Natural killer cells induce distinct modes of cancer cell death: Discrimination, quantification, and modulation of apoptosis, necrosis, and mixed forms. *J. Biol. Chem.* **293**, 16348–16363 (2018). doi: [10.1074/jbc.RA118.004549](#); pmid: [30190323](#)
- A. Zychlinsky, L. M. Zheng, C. C. Liu, J. D. Young, Cytolytic lymphocytes induce both apoptosis and necrosis in target cells. *J. Immunol.* **146**, 393–400 (1991). pmid: [1984450](#)
- D. K. Blanchard et al., Role of extracellular adenosine triphosphate in the cytotoxic T-lymphocyte-mediated lysis of antigen presenting cells. *Blood* **85**, 3173–3182 (1995). doi: [10.1182/blood.V85.11.3173.bloodjournal85113173](#); pmid: [7538819](#)
- N. J. Waterhouse et al., Cytotoxic T lymphocyte-induced killing in the absence of granzymes A and B is unique and distinct from both apoptosis and perforin-dependent lysis. *J. Cell Biol.* **173**, 133–144 (2006). doi: [10.1083/jcb.200510072](#); pmid: [16606695](#)
- P. A. Henkart, M. S. Williams, C. M. Zacharchuk, A. Sarin, Do CTL kill target cells by inducing apoptosis? *Semin. Immunol.* **9**, 135–144 (1997). doi: [10.1006/snim.1997.0063](#); pmid: [9194224](#)
- D. Chowdhury, J. Lieberman, Death by a thousand cuts: Granzyme pathways of programmed cell death. *Annu. Rev. Immunol.* **26**, 389–420 (2008). doi: [10.1146/annurev.immunol.26.021607.090404](#); pmid: [18304003](#)
- J. Lieberman, Granzyme A activates another way to die. *Immunol. Rev.* **235**, 93–104 (2010). doi: [10.1111/j.0105-2896.2010.00902.x](#); pmid: [20536557](#)
- K. Ebnét et al., Granzyme A-deficient mice retain potent cell-mediated cytotoxicity. *EMBO J.* **14**, 4230–4239 (1995). doi: [10.1002/j.1460-2075.1995.tb00097.x](#); pmid: [7556064](#)
- O. Susanto et al., Mouse granzyme A induces a novel death with writhing morphology that is mechanistically distinct from granzyme B-induced apoptosis. *Cell Death Differ.* **20**, 1183–1193 (2013). doi: [10.1038/cdd.2013.59](#); pmid: [23744295](#)
- T. Bergsbaken, S. L. Fink, B. T. Cookson, Pyroptosis: Host cell death and inflammation. *Nat. Rev. Microbiol.* **7**, 99–109 (2009). doi: [10.1038/nrmicro2070](#); pmid: [19148178](#)
- N. Kayagaki et al., Non-canonical inflammasome activation targets caspase-11. *Nature* **479**, 117–121 (2011). doi: [10.1038/nature10558](#); pmid: [22002608](#)
- J. Shi et al., Inflammasome caspases are innate immune receptors for intracellular LPS. *Nature* **514**, 187–192 (2014). doi: [10.1038/nature13683](#); pmid: [25119034](#)
- N. Kayagaki et al., Caspase-11 cleaves gasdermin D for non-canonical inflammasome signalling. *Nature* **526**, 666–671 (2015). doi: [10.1038/nature15541](#); pmid: [26375259](#)
- J. Shi et al., Cleavage of GSDMD by inflammatory caspases determines pyroptotic cell death. *Nature* **526**, 660–665 (2015). doi: [10.1038/nature15514](#); pmid: [26375003](#)
- J. Shi, W. Gao, F. Shao, Pyroptosis: Gasdermin-Mediated Programmed Necrotic Cell Death. *Trends Biochem. Sci.* **42**, 245–254 (2017). doi: [10.1016/j.tibs.2016.10.004](#); pmid: [27932073](#)
- L. Galluzzi et al., Molecular mechanisms of cell death: Recommendations of the Nomenclature Committee on Cell Death 2018. *Cell Death Differ.* **25**, 486–541 (2018). doi: [10.1038/s41418-017-0012-4](#); pmid: [29362479](#)
- R. A. Aglietti et al., GsdmD p30 elicited by caspase-11 during pyroptosis forms pores in membranes. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7858–7863 (2016). doi: [10.1073/pnas.1607769113](#); pmid: [27339137](#)
- J. Ding et al., Pore-forming activity and structural autoinhibition of the gasdermin family. *Nature* **535**, 111–116 (2016). doi: [10.1038/nature18590](#); pmid: [27281216](#)
- X. Liu et al., Inflammasome-activated gasdermin D causes pyroptosis by forming membrane pores. *Nature* **535**, 153–158 (2016). doi: [10.1038/nature18629](#); pmid: [27383986](#)
- L. Sborgi et al., GSDMD membrane pore formation constitutes the mechanism of pyroptotic cell death. *EMBO J.* **35**, 1766–1778 (2016). doi: [10.15252/embj.201694696](#); pmid: [27418190](#)
- P. Broz, P. Pelegrin, F. Shao, The gasdermins, a protein family executing cell death and inflammation. *Nat. Rev. Immunol.* **20**, 143–157 (2020). doi: [10.1038/s41577-019-0228-2](#); pmid: [31690840](#)
- C. Rogers et al., Cleavage of DFNA5 by caspase-3 during apoptosis mediates progression to secondary necrotic/pyroptotic cell death. *Nat. Commun.* **8**, 14128 (2017). doi: [10.1038/ncomms14128](#); pmid: [28045099](#)
- Y. Wang et al., Chemotherapy drugs induce pyroptosis through caspase-3 cleavage of a gasdermin. *Nature* **547**, 99–103 (2017). doi: [10.1038/nature22393](#); pmid: [28459430](#)
- K. Akino et al., Identification of DFNA5 as a target of epigenetic inactivation in gastric cancer. *Cancer Sci.* **98**, 88–95 (2007). doi: [10.1111/j.1349-7006.2006.00351.x](#); pmid: [17083569](#)
- M. S. Kim et al., Aberrant promoter methylation and tumor suppressive activity of the DFNA5 gene in colorectal carcinoma. *Oncogene* **27**, 3624–3634 (2008). doi: [10.1038/sj.onc.1211021](#); pmid: [18223688](#)
- K. Yokomizo et al., Methylation of the DFNA5 gene is frequently detected in colorectal cancer. *Anticancer Res.* **32**, 1319–1322 (2012). pmid: [22493364](#)
- C. J. Wang et al., The expression and regulation of DFNA5 in human hepatocellular carcinoma DFNA5 in hepatocellular carcinoma. *Mol. Biol. Rep.* **40**, 6525–6531 (2013). doi: [10.1007/s11033-013-2581-8](#); pmid: [24154762](#)
- B. Vanherberghen et al., Classification of human natural killer cells based on migration behavior and cytotoxic response. *Blood* **121**, 1326–1334 (2013). doi: [10.1182/blood-2012-06-439851](#); pmid: [23287857](#)
- R. Roger et al., CD28-mediated cytotoxicity of YT natural killer cells on B7-positive targets induces rapid necrotic death independent of granule exocytosis. *Cell. Immunol.* **168**, 24–32 (1996). doi: [10.1006/cimm.1996.0045](#); pmid: [8599836](#)
- S. Mahrus, C. S. Craik, Selective chemical functional probes of granzymes A and B reveal granzyme B is a major effector of natural killer cell-mediated lysis of target cells. *Chem. Biol.* **12**, 567–577 (2005). doi: [10.1016/j.chembiol.2005.03.006](#); pmid: [15911377](#)
- T. Zhao et al., Granzyme K cleaves the nucleosome assembly protein SET to induce single-stranded DNA nicks of target cells. *Cell Death Differ.* **14**, 489–499 (2007). doi: [10.1038/sj.cdd.4402040](#); pmid: [17008916](#)
- P. J. Beresford, C. M. Kam, J. C. Powers, J. Lieberman, Recombinant human granzyme A binds to two putative HLA-associated proteins and cleaves one of them. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9285–9290 (1997). doi: [10.1073/pnas.94.17.9285](#); pmid: [9256474](#)
- D. A. Thomas, C. Du, M. Xu, X. Wang, T. J. Ley, DFF45/ICAD can be directly processed by granzyme B during the induction of apoptosis. *Immunity* **12**, 621–632 (2000). doi: [10.1016/S1074-7613\(00\)80213-7](#); pmid: [10894162](#)
- C. L. Ewen, K. P. Kane, R. C. Bleackley, Granzyme H induces cell death primarily via a Bcl-2-sensitive mitochondrial cell death pathway that does not require direct Bcl-2 activation. *Mol. Immunol.* **54**, 309–318 (2013). doi: [10.1016/j.molimm.2012.12.020](#); pmid: [23352961](#)
- H. Lu et al., Granzyme M directly cleaves inhibitor of caspase-activated DNase (CAD) to unleash CAD leading to DNA fragmentation. *J. Immunol.* **177**, 1171–1178 (2006). doi: [10.4049/jimmunol.177.2.1171](#); pmid: [16818775](#)
- K. Sakuishi et al., Targeting Tim-3 and PD-1 pathways to reverse T cell exhaustion and restore anti-tumor immunity. *J. Exp. Med.* **207**, 2187–2194 (2010). doi: [10.1084/jem.20100643](#); pmid: [20819927](#)

47. S. I. Mosely *et al.*, Rational selection of syngeneic preclinical tumor models for immunotherapeutic drug discovery. *Cancer Immunol. Res.* **5**, 29–41 (2017). doi: [10.1158/2326-6066.CIR-16-0114](https://doi.org/10.1158/2326-6066.CIR-16-0114); pmid: [27923825](https://pubmed.ncbi.nlm.nih.gov/27923825/)
48. A. O. Kamphorst *et al.*, Rescue of exhausted CD8 T cells by PD-1-targeted therapies is CD28-dependent. *Science* **355**, 1423–1427 (2017). doi: [10.1126/science.aaf0683](https://doi.org/10.1126/science.aaf0683); pmid: [28280249](https://pubmed.ncbi.nlm.nih.gov/28280249/)
49. F. Arce Vargas *et al.*, Fc-optimized anti-CD25 depletes tumor-infiltrating regulatory T cells and synergizes with PD-1 blockade to eradicate established tumors. *Immunity* **46**, 577–586 (2017). doi: [10.1016/j.immuni.2017.03.013](https://doi.org/10.1016/j.immuni.2017.03.013); pmid: [28410988](https://pubmed.ncbi.nlm.nih.gov/28410988/)
50. N. Saeki *et al.*, Distinctive expression and function of four GSDM family genes (GSDMA-D) in normal and malignant upper gastrointestinal epithelium. *Genes Chromosomes Cancer* **48**, 261–271 (2009). doi: [10.1002/gcc.20636](https://doi.org/10.1002/gcc.20636); pmid: [19051310](https://pubmed.ncbi.nlm.nih.gov/19051310/)
51. Z. Tang, B. Kang, C. Li, T. Chen, Z. Zhang, GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **47** (W1), W556–W560 (2019). doi: [10.1093/nar/gkz430](https://doi.org/10.1093/nar/gkz430); pmid: [31114875](https://pubmed.ncbi.nlm.nih.gov/31114875/)
52. C. J. Froelich, J. Pardo, M. M. Simon, Granule-associated serine proteases: Granzymes might not just be killer proteases. *Trends Immunol.* **30**, 117–123 (2009). doi: [10.1016/j.it.2009.01.002](https://doi.org/10.1016/j.it.2009.01.002); pmid: [19217825](https://pubmed.ncbi.nlm.nih.gov/19217825/)

53. S. S. Metkar *et al.*, Human and mouse granzyme A induce a proinflammatory cytokine response. *Immunity* **29**, 720–733 (2008). doi: [10.1016/j.immuni.2008.08.014](https://doi.org/10.1016/j.immuni.2008.08.014); pmid: [18951048](https://pubmed.ncbi.nlm.nih.gov/18951048/)

ACKNOWLEDGMENTS

We thank Z. Shen for providing antibody to PD-1 and NK-92MI and A375 cells, Z. Dong for YTS cells, X. Zhao for SW837 and SKCO1 cells, W. Shen for Edman sequencing, Y. Sun for assistance with microscopy, and Y. Xu for assistances with mouse experiments. We also thank members of the Shao laboratory for their assistance and helpful discussions.

Funding: The work was supported by the Basic Science Center Project of the National Natural Science Foundation of China (81788104), the National Key Research and Development Program of China (2017YFA0505900 and 2016YFA0501500), the Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (2019-I2M-5-084), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB37030202), and a fund from Hoffmann-La Roche AG (ROADS027). **Author contributions:** Z.Z. and F.S. conceived the study; Z.Z., together with H.H. and Y.S., performed most of the experiments; K.W. contributed to the finding of GSDMB transcriptional priming; X.S., Yu.W., D.L., W.L., and J.D provided

technical supports and insights. Ya.W. and W.H. prepared CAR T cells; Y.Z. and Lia.S. provided TCR T cells; Lin.S. provided tumor tissue samples; and Z.Z. and F.S. analyzed the data and wrote the manuscript. F.S. obtained the funding and supervised the study. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. All the plasmids and cell lines generated in this study are available from the authors under a materials transfer agreement with the National Institute for Biological Sciences.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6494/eaaz7548/suppl/DC1

Materials and Methods

Figs. S1 to S11

References (54–59)

Movies S1 and S2

[View/request a protocol for this paper from Bio-protocol.](#)

7 October 2019; accepted 3 April 2020

Published online 16 April 2020

10.1126/science.aaz7548

RESEARCH ARTICLE SUMMARY

MAGNETISM

Self-induced spin glass state in elemental and crystalline neodymium

Umut Kamber, Anders Bergman, Andreas Eich, Diana Iuşan, Manuel Steinbrecher, Nadine Hauptmann, Lars Nordström, Mikhail I. Katsnelson, Daniel Wegner*, Olle Eriksson, Alexander A. Khajetoorians*

INTRODUCTION: Spin glasses are one of the more intriguing, but least understood, magnetic states of matter. In stark contrast to ordered magnets, spin glasses form a state characterized by seemingly random and uncorrelated magnetic patterns lacking long-range order. The distinguishing feature of spin glasses is aging: The magnetic state depends on its history, which is driven by multiple time scales. Despite decades of theoretical developments, there is still no clear understanding about when spin glass behavior arises. It is commonly believed that disorder is a key ingredient, in addition to competing magnetic interactions, as exhibited by the textbook example of dilute magnetic alloys.

RATIONALE: Despite more than 50 years of investigation, there is still no consensus on the magnetic ground state of the element neodymium (Nd). Below the Néel temperature, previous experiments reported the onset of static spin spirals with different periodicities, or so-called “multi- Q states,” as well as evidence of additional phase transitions. Although these observations illustrate the complexity in this material, it is not well understood how the multiplicity of these Q states depends on tem-

perature and the exchange landscape of the material and which real-space magnetic interactions cause the multi- Q states.

RESULTS: We show that single-crystalline elemental Nd exhibits unconventional spin glass behavior well described by the recently proposed concept of self-induced spin glassiness. Traditional spin glasses, such as the hallmark metallic alloy Cu-Mn, are based on competing interactions and disorder. By contrast, self-induced spin glasses can be created solely by competing interactions without strong disorder, for example, in systems with long-range magnetic interactions. We studied Nd by growing ultraclean, epitaxial thick islands and films, both of which were representative of bulk Nd. Using spin-polarized scanning tunneling microscopy, magnetic images of the surface revealed strong, local, noncollinear magnetic order at the atomic scale with a multiplicity of Q states and no long-range order (left). We quantified the role of defects on the glassiness, showing that cleaner samples led to more pronounced glassy behavior exemplified by the increased mixing of distinct Q states. The spatially dependent Q states were defined by

a spectral distribution of degenerate magnetic wave vectors (right), which varied spatially and with time. Harnessing ab initio and atomic spin dynamics calculations, we quantified the magnetic interactions, which illustrated strongly competing, distance-dependent interactions intricately linked to the crystal structure of

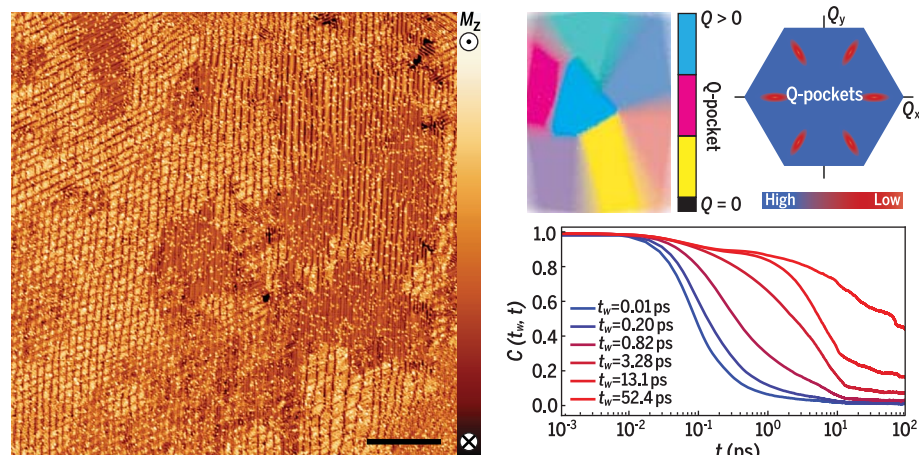
Nd. Moreover, the resultant energy landscape illustrates many favorable Q states, showing that glassiness in Nd results from the conditions required for self-induced spin glasses.

ON OUR WEBSITE

Read the full article at <https://dx.doi.org/10.1126/science.aay6757>

When probing the response to applied magnetic fields and temperature, we observed the existence of multiple time scales reminiscent of aging in traditional spin glass materials. The calculated autocorrelation function, a mean-field picture of the aging dynamics, also showed aging behavior reminiscent of traditional spin glasses (bottom). Experiments also showed that the aging dynamics were strongly dependent on the underlying value of Q , leading to behavior reminiscent of the concept of dynamic heterogeneity observed in structural glasses. In this way, the energy landscape can be described by a rugged, multiwell potential composed of degenerate Q states. Unlike traditional spin glasses, competing interactions in Nd are driven by its electronic and the structural properties, which lead to self-induced glassiness.

CONCLUSION: Our findings not only suggest that glassy behavior can be found in elements with crystalline order, but also unravel the decades-long debate about the magnetic behavior of Nd. The coexistence of short-range order exhibited by degenerate Q states and aging dynamics manifested by multiple time scales provides the first experimental confirmation of self-induced glassiness. The existence of strong local Q order and the Q -dependent dynamics may imply that multiple but select time scales exist, resulting from a rugged, multiwell landscape, when compared with traditional spin glasses. This provides a new platform with which to explore dynamic heterogeneity in a spin glass material. It remains to be understood what is particularly special about the crystal structure and electronic properties in Nd, and this will necessitate a deeper theoretical understanding of the role of electron correlation effects, as well as the interplay between spin and orbital degrees of freedom. The example here expands our views on magnetic states of matter, inviting further research into aging behavior in magnetic systems. ■



Spin- Q glass. (Left) Real-space magnetization image with spin-polarized scanning tunneling microscopy at $T = 1.3$ K of thick films of Nd(0001). The surface shows multi- Q states but no long-range order. (Right) Sketch of spin- Q glass in both real and reciprocal spaces, with color illustrating the distribution of Q states in real space derived from flat pockets in Q -space. (Bottom) Calculated autocorrelation function for Nd with increasing waiting time (t_w) illustrating aging behavior.

The list of author affiliations is available in the full article online.

*Corresponding author. Email: d.wegner@science.ru.nl (D.W.); a.khajetoorians@science.ru.nl (A.A.K.)

Cite this article as U. Kamber *et al.*, *Science* **368**, eaay6757 (2020). DOI: [10.1126/science.aay6757](https://doi.org/10.1126/science.aay6757)

RESEARCH ARTICLE

MAGNETISM

Self-induced spin glass state in elemental and crystalline neodymium

Umut Kamber¹, Anders Bergman², Andreas Eich¹, Diana Iuşan², Manuel Steinbrecher¹, Nadine Hauptmann¹, Lars Nordström², Mikhail I. Katsnelson¹, Daniel Wegner^{1*}, Olle Eriksson^{2,3}, Alexander A. Khajetoorians^{1*}

Spin glasses are a highly complex magnetic state of matter intricately linked to spin frustration and structural disorder. They exhibit no long-range order and exude aging phenomena, distinguishing them from quantum spin liquids. We report a previously unknown type of spin glass state, the spin- Q glass, observable in bulk-like crystalline metallic neodymium thick films. Using spin-polarized scanning tunneling microscopy combined with *ab initio* calculations and atomistic spin-dynamics simulations, we visualized the variations in atomic-scale noncollinear order and its response to magnetic field and temperature. We quantified the aging phenomena relating the glassiness to crystalline symmetry and the energy landscape. This result not only resolves the long-standing debate of the magnetism of neodymium, but also suggests that glassiness may arise in other magnetic solids lacking extrinsic disorder.

Spin glasses are one of the more intriguing, but least understood, magnetic states of matter (1–5). Ferromagnets or antiferromagnets form a long-range ordered state when cooled but spin glasses form a state characterized by seemingly random and uncorrelated magnetic patterns. The magnetization pattern in spin glasses can be compared to the amorphous structure of glasses such as silicon dioxide, which exhibit local structural correlations but lack a long-range ordered state. Interest in spin glasses spans many fields, ranging from iron-based superconductors (6) to theoretical machine learning (4, 5, 7), and they have been suggested to be relevant in quantum topological excitations.

Spin glasses are characterized by a glass transition temperature and by aging; i.e., the magnetic state depends on its history, which is driven by a distribution of distinctive spin-relaxation processes with time scales spanning many orders of magnitude (4, 5). Aging also distinguishes spin glasses from so-called quantum spin liquids (8) or spin ices (9), which remain disordered down to zero temperature because of quantum fluctuations and lack of memory. The paradigm of all these types of complex magnets involves magnetic frustration derived from geometry or competing interactions. However, unlike spin liquids, disorder is traditionally also considered necessary to drive non-ergodic behavior in spin glasses.

There is still no clear understanding about when spin glass behavior can arise in mag-

netic materials. The most commonly debated models (1, 3–5) describe randomly distributed spins with long-range magnetic interactions of alternating ferromagnetic and antiferromagnetic coupling, such as that seen in prototypical materials such as dilute magnetic alloys (1, 10). In the thermodynamic limit, spin glasses have a hierarchical energy landscape with infinitely many local energy minima separated by energy barriers of multiple heights; therefore, there is a broad distribution of transition times between different minima (3–5, 11) that results in an absence of local and long-range order. Most models of spin glasses invoke structural disorder, i.e., amorphousness, as a key requirement along with competing magnetic interactions (12). However, self-induced glassiness, a concept introduced initially for the stripe glass behavior in high-temperature superconductors (13, 14), was later developed for magnets (15, 16). Within this framework, competing interactions alone can lead to the glassy state, even in the absence of external disorder, and can give rise to intermediate regimes that exhibit multiwell potentials (17).

We show that single-crystalline elemental neodymium (Nd) exhibits a previously unknown type of spin glass behavior. Spin-polarized scanning tunneling microscopy (SP-STM) on the (0001) surface of thick Nd films revealed that the magnetic state exhibited strong local noncollinear magnetic order but lacked a long-range ordered state. This local order is defined by a spectral distribution of degenerate magnetic wave vectors, or Q states, which varied spatially and with time. We probed the response of this so-called spin- Q glass to applied magnetic fields and variable temperature to quantify its aging behavior and energy landscape. Harnessing *ab initio* methods and sim-

ulations, we quantified the competing long-range magnetic interactions and the favorable Q states, illustrating that this unconventional glassy behavior results from valley-like pockets of degenerate Q states, as has been proposed for self-induced spin glasses (15, 16). Moreover, we performed calculations of the autocorrelation function for pristine Nd and also found multiple relaxation times in its spin dynamics. Our findings not only suggest that glassy behavior and aging can be found in systems with crystalline order, but also unravel an unresolved debate about the magnetic ground state of elemental Nd that has challenged scientists for several decades.

Magnetic ground state of Nd(0001)

Despite more than 50 years of investigations, there is still no consensus on the magnetic ground state of Nd and the origin of the complexity reported in various experimental observations (18–26). Below the Néel temperature (T_N), neutron diffraction allowed observation of the onset of static multi- Q states with decreasing temperature. However, it is not well understood how the multiplicity of these Q states depends on the exchange landscape of the material and which real-space magnetic interactions cause these various Q states. Moreover, other measurements indicated additional phase transitions below T_N , suggesting that the original conclusions from neutron diffraction of a modulated antiferromagnetic structure were oversimplified (23). In addition, there are experimental observations above T_N that provide evidence for short-range order as well as a strongly frustrated exchange landscape (26, 27). These open questions illustrate the need for a characterization of the exchange interactions in Nd, as well as a real-space characterization of the local magnetic order.

Atomic-scale visualization of the spin- Q glass state

To characterize the atomic-scale magnetization of the surface of Nd(0001), we used SP-STM and spin-polarized scanning tunneling spectroscopy (SP-STs) (28). We epitaxially grew thick films [up to 100 monolayers (ML)] of Nd(0001)/W(110). We note that lanthanide films prepared in this way on various bcc(110) substrates exhibit high crystallinity and superior surface cleanliness over sputter-annealed bulk single crystals (29–37). It has been shown for various lanthanide elements, including Nd, that films grown with bulk lattice parameters above a thickness of 10 ML (31, 35, 38–40) exhibit a bulk-like electronic structure >30 ML (32, 33, 41) and start to exhibit bulk-like magnetic behavior in the range of 30 to 50 ML (42–45). Superlattice studies showed that 33 to 39 ML-thick Nd films already exhibit the temperature-dependent neutron scattering features known from bulk single crystals (38, 46, 47).

¹Institute for Molecules and Materials, Radboud University, Nijmegen, Netherlands. ²Department of Physics and Astronomy, Uppsala University, Uppsala, Sweden. ³School of Science and Technology, Örebro University, Örebro, Sweden. *Corresponding author. Email: d.wegner@science.ru.nl (D.W.); a.khajetoorians@science.ru.nl (A.A.K.)

The sample morphology can be tuned to either layer-by-layer grown closed films or islands [Stranski-Krastanov (SK) growth], depending on the annealing temperature (39, 42, 48–50). We grew two types of samples: SK-grown islands of >50 ML thickness (Fig. 1, A and B) and closed, ~100 ML films (fig. S1B) (50). As discussed above, SK islands should readily represent bulk-like structural, electronic, and magnetic properties. We verified this on thicker (~100 ML) closed films, which showed identical magnetization patterns, as we discuss in the supplementary text, section S2 (50). However, we observed that closed thick films show inferior surface qualities because of the presence of more impurities, as well as screw dislocations that can lead to pinning of magnetic structures (51). Because the latter was not present for our island-grown samples (fig. S1A), we focused on data taken from islands with thickness between 58 and 92 ML and lateral sizes between 58,000 and 200,000 nm², which may account for differences in previous measurements. Further details of the sample preparation and morphology, as well as a discussion of thin films versus bulk samples, can be found in the supplementary text, section S1 (50).

Using STS at low temperature (0.03 to 7 K), we probed the Nd(0001) surface state, the presence and sharpness of which has been shown to be a probe of the cleanliness of the film (27, 31, 32, 52, 53). The surface state is characterized by an exchange splitting into a

majority peak visible below the Fermi energy (E_F) [sample bias (V_S) = 0] and a minority peak above E_F , with an additional narrow peak at E_F (Fig. 1C) (49). We saw no difference between spectra taken on the islands and those taken on locally clean areas of the thick film. Furthermore, the magnetic exchange splitting reproduced previous low-temperature data taken on a 30 ML film (27, 54), which is further evidence that our samples are beyond the thin-film limit and fully reflect bulk electronic and magnetic properties.

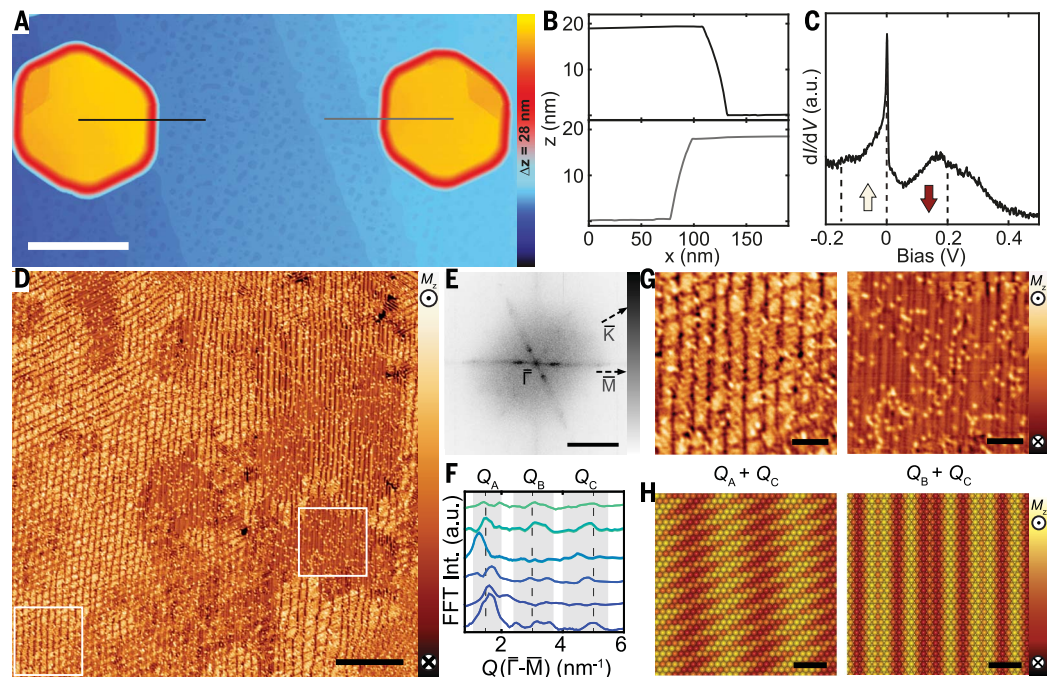
To get magnetic contrast, we used the spin-polarized nature of the exchange-split surface state and imaged the surface in constant-current mode at two characteristic voltages representing dominant tunneling into the minority state (V_S = 200 mV) and out of the majority state (V_S = -150 mV), respectively (fig. S2) (50). We used an out-of-plane sensitive antiferromagnetic chromium (Cr) probe, which relates spin-dependent contrast variations directly to the z -projection, the c -axis in the double hexagonally close packed (dhcp) structure, of the magnetization. We consider the subtracted image to remove stronger topographic modulations resulting from buried substrate steps or locally varying sample thickness. We refer to this subtracted image as the magnetization image (Fig. 1, D and G, and fig. S2D). More details on this image-processing procedure can be found in the supplementary text, section S2 (50).

The magnetization images of the surface revealed strong and clear short-range magnetic order with periodicities λ varying from 0.9 to 4.5 nm and oriented along or near the high-symmetry axes. These atomic-scale variations are directly related to local noncollinear magnetic order, with varying periodicities depending on spatial location, which is defined by a superposition of local magnetic wave vectors $Q_i = 2\pi/\lambda_i$ (Fig. 1, G and H, and supplementary text, section S3) (50). Unexpectedly, although clear short-range order could be seen, defined by a local multi- Q state, there was no observable long-range ordered state found for any of the probed experimental conditions.

To better visualize the variations of local multi- Q order, we considered reciprocal-space images obtained through fast Fourier transform (FFT) of the real-space magnetization images, which we refer to as Q -space images. A signature of the lack of long-range order and competing short-range order can be directly visualized by the smeared and broadly distributed spectral weight in various regions in Q -space (Fig. 1E). Q vectors here were derived from real-space SP-STM maps and were not directly measured as in the case of neutron diffraction (see supplementary text, section S3, for further discussion) (50). We only produced and analyzed Q -space images from within a flat terrace on individual islands (typical width, ~150 nm). Thus, the measured broadening

Fig. 1. Imaging the spin- Q glass state of Nd(0001).

(A) Constant-current STM image of an SK-grown Nd film on W(110) revealing nearly flat-top islands on an Nd wetting layer. The surface defect concentration is <0.01 ML [scale bar, 150 nm; V_S = 1 V; tunneling current (I_t) = 20 pA]. (B) Line profiles along the indicated island edges. The thickness of each island is >50 ML. (C) dI/dV spectrum acquired on the Nd(0001) surface showing the exchange-split surface state [stabilization bias (V_{stab}) = 1 V; stabilization current (I_{stab}) = 200 pA; modulation bias (V_{mod}) = 1 mV; T = 40 mK]. (D) Magnetization image illustrating the spatially complex magnetic ground state of a spin- Q glass, which lacks long-range order (T = 1.3 K; B = 0 T; I_t = 200 pA). The contrast is directly related to variations in the out-of-plane magnetization (M_z) imaged with an out-of-plane sensitive Cr bulk tip (scale bar, 50 nm). (E) Q -space image of the magnetization image in (D) (scale bar, 5 nm⁻¹) illustrating a large distribution of states in Q -space. (F) Line cuts along $\bar{\Gamma}$ to $\bar{M}$ of various Q -space images taken from smaller



sections of the image in (D) (compare fig. S5). (G) Close-up views of regions marked in (D) illustrating the local spatial variation of the magnetic order (scale bar, 10 nm). (H) Schematic of the out-of-plane projected magnetization resulting from a superposition of the labeled Q vectors (scale bar, 2 nm). The wave vector amplitudes used are marked with dashed lines in (F).

resulted from the spectral distribution of Q , not from averaging over different islands with different orientations. To highlight this, Q -space images were also produced from smaller spatial regions of the same image, which illustrated sharpened and characteristic spectral weight of Q vectors compared with the larger-scale image (Fig. 1G and fig. S4). Various line cuts along the Γ to M direction of multiple Q -space images taken from different spatial regions of Fig. 1D are illustrated in Fig. 1F. The resultant plots revealed a spatial variation in spectral weight in at least three distinct regions, or Q -pockets, with substantial spectral weight along the high-symmetry axes.

In the ensuing discussion, we focus particularly on the three Q -pockets with wave vectors at $Q_A = 1.1$ to 2.0 nm^{-1} , $Q_B = 2.7$ to 3.5 nm^{-1} , and $Q_C = 4.6$ to 5.3 nm^{-1} . From this information, we plotted a schematic of the out-of-plane projected magnetization with respect to the atomic lattice (Fig. 1H) for the two regions shown in Fig. 1G with the defined Q vectors from these Q -pockets. The corresponding Q -space images are a direct visualization of the multi- Q nature acquired over the spatial area of the given images, and they provide a quantitative comparison to previous neutron diffraction studies (18–21, 24–26). The spectral weight around these pockets, as well as blurring of the intensity of the Q states in larger-scale images (Fig. 1E), was an initial indication of the glassy nature of the magnetic state and reminiscent of spin-based analogs of stripe or checkerboard order in strongly correlated compounds (55).

To illuminate the concept of a spin- Q glass, we qualitatively illustrate the energy landscape in Q -space images for a spin- Q glass compared

with a ferromagnet (Fig. 2). A long-range ordered state can be related to a global minimum in Q -space (28, 56, 57), where a single-domain ferromagnetic state is equivalent to a $Q = 0$ global minimum (Fig. 2A). By contrast, a spin- Q glass is distinguished by the existence of flat valleys defined by a distribution of many local minima, i.e., Q -pockets, at finite Q values (Fig. 2B). Broad Q -pockets led to a lack of a preferentially long-range ordered state. Instead, there were local regions defined by a strong local order derived from a spectral weight of mixed Q vectors within the given pockets, and different regions exhibited random distributions of this spectral weight (see color image in Fig. 2B). The superposition of this spatially varying magnetization led to an overall broadening in the spectral weight. Note that multi- Q states for thin $3d$ transition metal films show well-ordered domains (58). Within the concept of self-induced glassiness for spins, such pockets may result from a strong competition of magnetic interactions, leading to highly degenerate states (15, 16).

Theoretical analysis of the magnetic landscape of Nd(0001)

To analyze the origin of the spin- Q glass state in Nd(0001) and its unexpected magnetic patterns (59), we used ab initio calculations to quantify and understand the exchange interactions and the energy landscape in Q -space of bulk Nd, which adopts a dhcp structure that is critical for its magnetic exchange interactions (Fig. 3A). The RSPt code was used for this purpose (see the materials and methods and supplementary text, section S4) (50). The calculated exchange interactions for bulk Nd are shown for both the dhcp and a hypothetical hcp

structure. There is a minute energy difference between these crystal structures that results from the large similarities in atomic arrangement.

The distance dependence of the exchange interactions in the hcp structure illustrates a prototypical behavior with ferromagnetic nearest-neighbor interactions and an oscillating Ruderman-Kittel-Kasuya-Yosida (RKKY)-like interaction at large distances, as seen with other lanthanides such as gadolinium (Gd) (60, 61). By contrast, in the dhcp, at shorter range, the interactions are much weaker and primarily antiferromagnetic, but at larger distances, an RKKY interaction sets in. These strong competing magnetic interactions in Nd create conditions for frustrated magnetism and spin glass formation, as described for self-induced glassiness (15–17). Calculated values of the magnetic moment of the surface atoms, and for deeper layers, were similar to the bulk value. The bulk moment was restored within three layers beneath the surface (see supplementary text, section S4) (50).

To clarify the impact of the calculated exchange interactions in the dhcp structure of Nd, we evaluated the magnetic energy landscape by means of single- Q spin spirals, magnetic structures that can be parametrized by a single wave vector. For this purpose, we calculated the energy of helical spin spirals. Using the calculated magnetic exchange interactions from Fig. 3A, we parametrized an effective Heisenberg spin Hamiltonian from which the energy $E(\mathbf{Q})$ of the single- Q spirals was then calculated (Fig. 3B). The energy-landscape exploration was performed by fixing the z -component Q_z and then sweeping over Q_x and Q_y in the first Brillouin zone (BZ) (see supplementary text, section S6) (50).

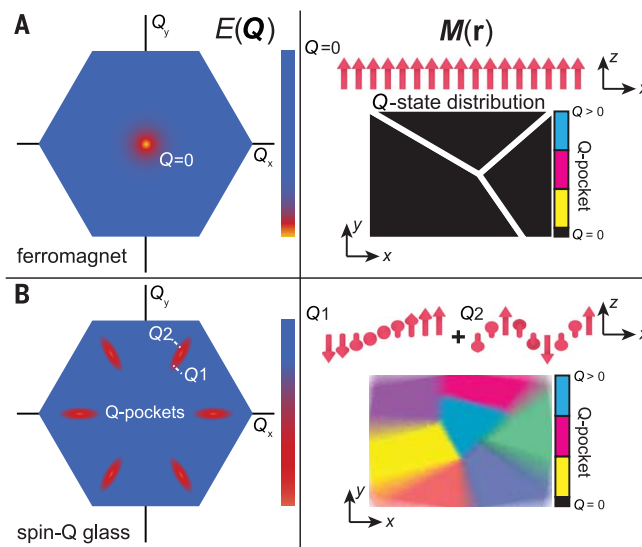
In Fig. 3B, we present the single- Q energy landscape for all possible Q_x and Q_y combinations in the cell spanned by the dhcp reciprocal lattice vectors for the case when $Q_z = 2\pi/c$, which corresponded to the configuration with the lowest single- Q energy. This choice of Q_z corresponded to a 90° rotation of the moments in adjacent atomic layers within the dhcp unit cell. The Q -dependent energy is color-coded such that red (blue) regions correspond to spin spirals with low (high) energy. This visualization illustrates the complex energy landscape of Nd, as exemplified by the broad and flat, dark-red ring-like structure similar to the deduced landscape in (26).

Instead of a distinct, sixfold degenerate set of strong energy minima that could be expected for a spin-spiral magnet on a hexagonal lattice, the red ring structure showed that the energy barriers between global and local minima in this region of the BZ were very small. In addition, high-energy local minima, or pockets (in red), were distributed with hexagonal symmetry just inside the BZ, and another low-energy valley-like structure formed along

Fig. 2. Energy landscape of a spin- Q glass.

(A) Q -space image of the energy landscape $E(\mathbf{Q})$ of a prototypical ferromagnet, which exhibits a strong global minimum at $Q = 0$, corresponding to a real-space magnetization pattern $\mathbf{M}(\mathbf{r})$ where all spins are aligned. Further minimization leads to the formation of distinct domains (black) separated by domain walls (white), but all domains are defined by a repeating Q -state distribution, where $Q = 0$.

(B) Characteristic Q -space image for a spin- Q glass, which can be distinguished by flat valleys (Q -pockets) at nonzero Q values, which leads to a superposition of a distribution of Q states with different periodicities residing in each pocket. This results in a complex $\mathbf{M}(\mathbf{r})$ pattern that lacks long-range order. The spatial distribution of Q states contains regions with local order defined by mixing of Q states (colors) derived from the given Q -pockets.



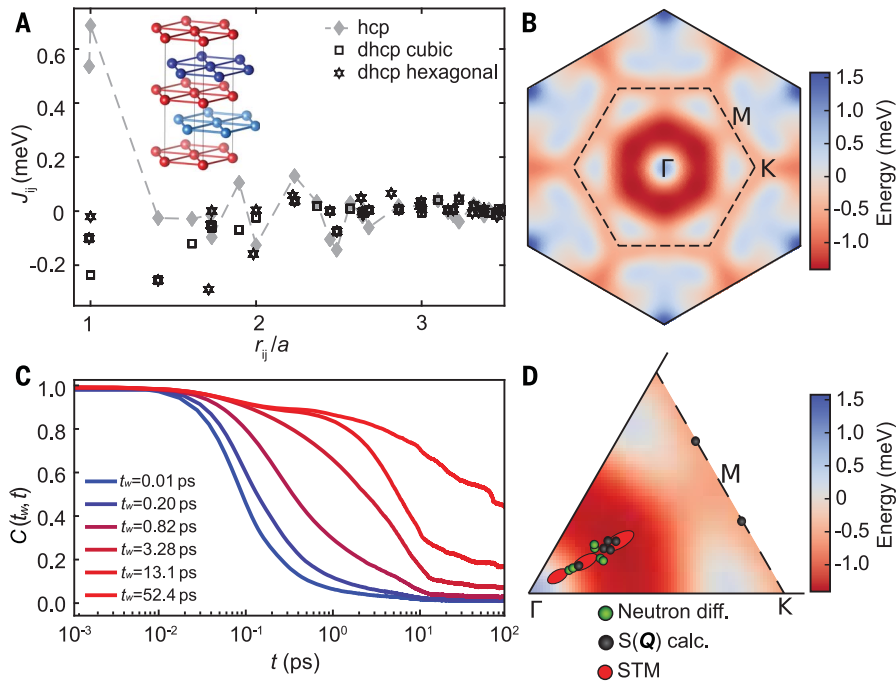


Fig. 3. Elemental Nd electronic and magnetic landscape. (A) Calculated Heisenberg magnetic exchange interactions among Nd spin moments with magnitude $2.454 \mu_B$ in both the dhcp Nd (black) structure and the hypothetical hcp structure (gray). A negative interaction denotes a preference for an antiferromagnetic alignment among the spins, and a positive one denotes a ferromagnetic alignment. Inset, dhcp crystal structure with an ABAC stacking, where the cubic A sites are represented by red spheres and the hexagonal B and C sites by light and dark blue spheres, respectively. (B) Energy landscape for single- Q spin spirals with $\mathbf{Q} = (Q_x, Q_y, 2\pi/c)$ as evaluated from the calculated exchange interactions for the dhcp structure. (C) Autocorrelation function $C(t_w, t) = \langle \mathbf{m}_i(t + t_w) \cdot \mathbf{m}_i(t_w) \rangle$ for dhcp Nd at $T = 1$ K. (D) Comparison of Q states for SP-STM (this work, red), neutron diffraction [(21, 25), green], and simulations (this work, black).

the BZ boundary. Magnetic anisotropy was not considered here and we expect that it would bias the ring structure toward the high-symmetry directions, creating pockets akin to those seen in the experiments. Thus, Fig. 3B shows that the energy landscape of Nd has several broad Q -pockets, which supports the formation of the experimentally observed spin- Q glass structure. The presence of strongly competing interactions leading to a glass-like energy landscape in Q -space is a key manifestation of the concept of self-induced glassiness (13–15).

Spin- Q glass dynamics: autocorrelation and static correlation function

In addition to the single- Q energy-landscape explorations, we also used Monte Carlo and atomistic spin dynamics (ASD) simulations to find the ground-state magnetic structure. Although not all conventional spin glasses exhibit identical relaxation dynamics, a common trait is aging, i.e., that the relaxation process slows down over time and never settles on a single equilibrium state. Aging is observed in the Edwards–Anderson model (1). To characterize the aging dynamics of dhcp Nd, we studied the two-time autocorrelation function

defined as $C(t_w, t) = \langle \mathbf{m}_i(t + t_w) \cdot \mathbf{m}_i(t_w) \rangle$, where the brackets denote averaging over all sites of the system, $\mathbf{m}_i$ is the atomic moment, and t_w is waiting time. If a system relaxes to a fixed ground state following nonglassy dynamics, then the autocorrelation function should increase with increasing t_w . Autocorrelation analysis based on ASD has previously been used to capture the multiple relaxation scales of a conventional spin glass system (10). This approach, which is typically analyzed with a mean-field approach, is well suited to characterizing the multiple time scales indicative of aging (see supplementary text, section S8) (50).

In Fig. 3C, we show the autocorrelation at logarithmically spaced t_w for dhcp Nd as it relaxed from a fully disordered state at $T = 1$ K. The autocorrelation function decayed exponentially toward zero, a typical feature of aging and spin glass behavior (Fig. 3C) (50) and was similar to that of a traditional spin glass system such as Cu–Mn or that of the Edwards–Anderson model (see supplementary text, section S8) (10, 50). In other words, the relaxation process of Nd never arrived at a well-defined energy minimum. We note the autocorrelation analysis often assumes a mean-

field picture of aging, whereas a more complete picture should include the Q -dependent contributions to the multiple time scales. These simulations were performed for bulk dhcp Nd, meaning that the observed spin- Q glassy behavior can only be a result of the exchange interactions of the system because neither surface states nor defects were present in the simulations.

In addition to mapping out the aging dynamics of the spin- Q glass state, from these simulations, we obtained real-space spin structures (see supplementary text, section S7) (50) as well as the static correlation function $S(\mathbf{Q})$, which can be compared directly with the experimental Q -space images. In Fig. 3D, we present the simulated peaks of $S(\mathbf{Q})$ in comparison with (i) the observed range of spectral weight from the Q -pockets Q_A , Q_B , and Q_C in the SP-STM experiments [applied magnetic field (B_z) = 0 for both pristine zero-field cooled samples and samples after being exposed to magnetic field] and (ii) reported neutron diffraction data (21, 25). By comparison, the SP-STM and neutron diffraction data agreed well with the simulated results in regions of lower Q values. This agreement indicated that the surface-derived measurements were strongly coupled to the bulk magnetic properties. The simulations also provide distinct local maxima of the correlation function, which so far have not been detected experimentally.

Impact of impurities on Q -state distribution

Nd films illustrate characteristics of bulk Nd with the dhcp structure, including the expected surface state of Nd(0001) and the expected strain-free lattice constant (50). Although we did not observe bulk screw dislocations and large-scale defects for the island samples, the surface did show the presence of surface impurities. The presence of the interface did not seem to influence the imaged magnetism, e.g., underlying substrate step edges (see supplementary text, section S2) (50). Before discussing the role of the impurities on the magnetic order, we note that the main concentration of impurities in the source material is oxygen and carbon, each about 2000 parts per million by atom (table S1) (50). The impurity density of our surfaces is ~ 0.01 ML, which is below the detection limit of any surface-averaging spectroscopic technique and among the best reported for lanthanide metal surfaces (53, 62). Of these impurities, we observed the unperturbed intrinsic width of the surface state (54). Because the quality of the surface state (i.e., its intensity and width) in spatially averaging photoemission has been demonstrated to illustrate the structural order as well as cleanliness of the surface (31, 32, 52), we conclude that the surface defects in the present limit have a negligible impact on the overall electronic structure.

In Fig. 4, we show a comparison between two separately prepared samples and their respective FFTs, for which the defect density is sufficiently different. For samples with higher defect densities, imaged at $T = 1.3$ K, we observed more well-defined Q vectors at higher overall Q values, with a narrower distribution in Q -space compared with samples where the defect density was lower (e.g., Fig. 1). Although local order in this case had a well-defined periodicity, the orientation was not perfectly aligned, leading to a small, angular distribution of Q states. In stark contrast, as the defect density was reduced, the distribution of Q states strongly broadened. Thus, with cleaner samples, there was a stronger manifestation of multiple energy minima as more Q states stabilized. This observation contrasts with typical magnetically ordered systems, where any collinear or noncollinear configuration narrows its distribution in reciprocal space with cleaner samples. In this context, extended line defects (screw and edge dislocations) in lanthanide films were shown to pin magnetic order, whereas local point defects such as atomic adsorbates did not (57). As shown in figs. S1 and S3 (50), thick closed Nd films exhibit line defects as well as higher surface defect densities than the typical islands shown in Fig. 1. The unavoidable presence of dislocations and the higher amount of impurities in bulk single crystals may lead to defect-induced pinning of Q vectors, exemplifying the importance of cleanliness to observe the glassy behavior. However, for the ~ 100 ML-thick closed films, we observed qualitatively the same magnetic multi- Q structure as the islands (fig. S3) (50). Therefore, despite their seemingly small volume, the islands reflect the magnetic properties of bulk Nd. We conclude that below a critical defect density, a multiwell landscape emerged as a spin- Q glass phase, as illustrated in the defect-free calculations of both the energy landscape and the autocorrelation function (Fig. 3C).

Magnetic field evolution of the spin- Q glass state

We experimentally characterized the response of the magnetization to external magnetic fields. Out-of-plane field dependence is illustrated in Fig. 5 for a few chosen fields up to $B_z = 7$ T at $T = 1.3$ K (see supplementary text, section S9) (50). The application of magnetic field should favor states with smaller Q , eventually leading to a preferential Q state near the zone center (56, 63), when the Zeeman energy exceeds the local exchange energy. However, magnetic imaging of a given area at variable magnetic field revealed a different picture. At increasing magnetic fields, no distinct Q state became favorable, with the spectral weight strongly broadening along high-symmetry directions toward higher Q (Fig. 5G). The magnetic order was sensitive to fields on the order of $B_z = 0.5$ T, illustrating the degeneracy driven by

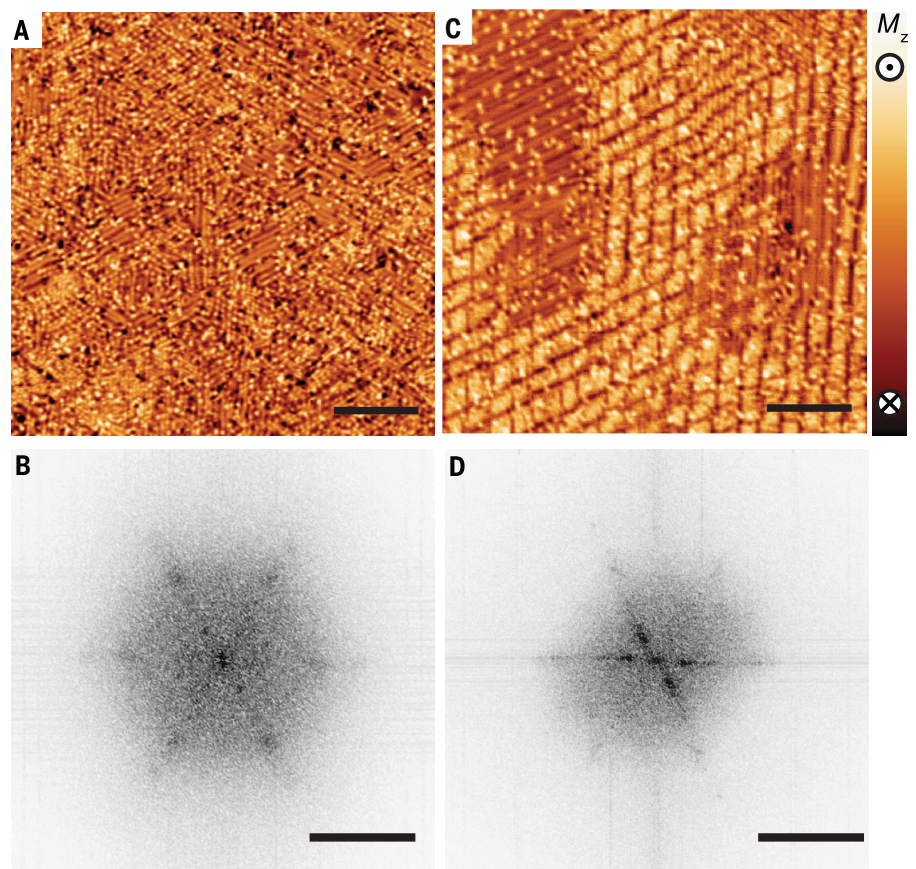


Fig. 4. Effect of defects on the spin- Q glass state. Magnetization and the corresponding Q -space images of a dirty surface (surface defect concentration of 0.03 ML) (A and B) and a clean surface (<0.01 ML) (C and D) (scale bar, 20 nm, $I_t = 200$ pA for magnetization images; scale bar, 5 nm $^{-1}$ for Q -space images, inverted grayscale). A higher amount of contamination results in pinning of the Q state around the defects, as illustrated in real-space magnetization images. Overall, measurements of several samples showed that the Q -state distribution is essentially unaltered for concentrations ≤ 0.014 ML. The Q -space image of the dirty sample shows more well-defined Q -pockets caused by the pinning, in contrast to the smeared-out Q -pockets along the high-symmetry axes in the clean sample.

the Q -pockets, whereas at the highest applied fields ($B_z = 7$ T), there was no distinct Q state, demonstrating the strong local exchange energy.

The application of in-plane fields revealed similar behavior: Favorable Q states collapsed onto an axis related to the direction of the applied field, but with appreciable smearing of the spectral weight along one particular axis that is reminiscent of the so-called archipelago phase seen in neutron scattering (22–24) (fig. S13) (50). The field-dependent behavior seen here cannot be attributed to the picture of local domains because there were neither clear domain boundaries that could be traced in the magnetic field nor a repeating or favorable Q structure. In addition, magnetostriction effects should saturate beyond 1 T at the measured temperature, as previously reported (23).

Aging of the spin- Q glass state through the magnetic field

Although field-dependent imaging illustrated the degeneracy within the Q -pockets, the dis-

tinguishing property of a spin glass is the observation of aging. Aging, in a mean-field picture, can be described by the existence of multiple relaxation time scales, as exemplified in Fig. 3C, leading to a magnetization state that never fully relaxes (10). To monitor the evolution of the ground state, we repeatedly ($i = 1, \dots, n$) applied the following procedure, starting from the pristine ($i = 0$) state: (i) sweep up the magnetic field to B_{zi} (maximum, 7 T) and stay at that value for a finite time τ_i , (ii) reduce the magnetic field to zero, and (iii) image the i^{th} zero-field state of the same area as the pristine state. Comparing several field-dependent cycles, the zero-field magnetic state at $T = 1.3$ K illustrates sequential redistributions of the spectral weight within and between the various Q -pockets (Fig. 6).

Consistent with aging, the magnetic state did not relax to a given ground state after perturbation, nor was there a clear tendency toward a different but particular ground state distribution. Instead, the overall trend was a

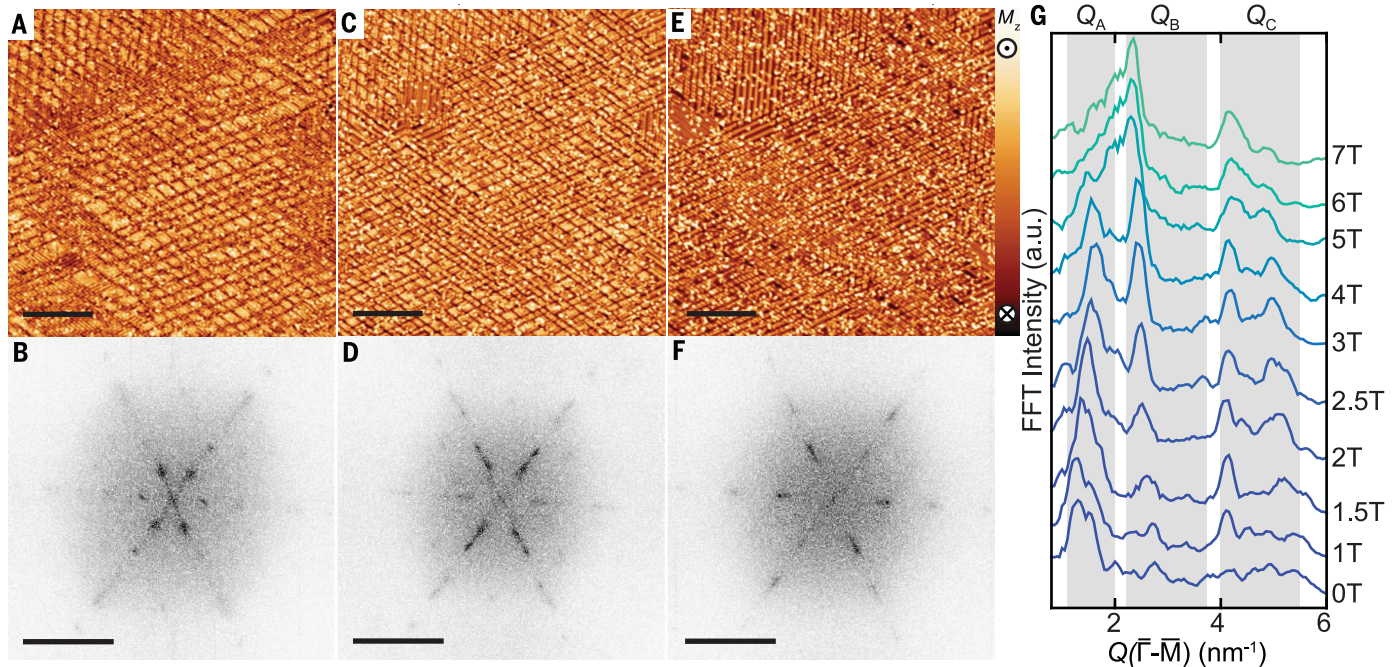


Fig. 5. Magnetic field evolution of the spin-Q glass state. Shown are magnetization and the corresponding Q-space images of the same area measured at $T = 1.3$ K in $B_z = 0$ T (**A** and **B**), in $B_z = 4$ T (**C** and **D**), and in $B_z = 7$ T (**E** and **F**) (scale bar, 30 nm, $I_t = 200$ pA for magnetization images; scale bar, 4 nm^{-1} for Q-space images, inverted grayscale). Increasing magnetic field does not favor a low Q state

nor exhibit the favorability of any Q state, as shown by the broadening of the spectral weight in the various pockets. (**G**) Line cuts along $\bar{\Gamma}$ to $\bar{M}$ of Q-space images (fig. S11) with finer intervals of increasing B_z . The spectral distribution becomes smeared out within the given pockets, leading to no well-defined long-range periodicity at the highest B_z . The surface defect concentration is 0.01 ML.

redistribution between all pockets toward a broader overall distribution along the high-symmetry directions, with the angular distribution sharpening along the high-symmetry axes. Moreover, the system never reverted to the initial zero-field cooled state. These findings ruled out that field-dependent cycling can be understood as the evolution of a favorable domain. To rule out hysteretic effects, we also performed field sweeps at positive and negative fields, but there was no clear correlation between such subsequent conditions, nor was there any correlation with local defects on the surface. Similar aging effects, and magnetic field evolution, were seen at different temperatures (fig. S15) (50), as we detail below, as well as with the application of an in-plane field and near island edges (fig. S14 and supplementary text, sections S10 and S11) (50).

Multiple and Q-dependent relaxation times at elevated temperature

The evolution of intermittent spectral weight distributions, which were frozen after subsequent field-sweep cycles, is a signature of slow aging dynamics and a hierarchy of states separated by small energies. By contrast, the application of temperature illustrated the presence of fast dynamics concomitant with slow relaxation dynamics. To illustrate this effect, we show the effect of temperature by imaging the same area at different temperatures (Fig. 7).

Compared with $T = 1.3$ K (Fig. 7B), the spectral weight within the Q_A pockets “melted” away at $T = 4.2$ K (Fig. 7D). In addition, in the real-space image, the associated long-wavelength pattern was no longer visible (Fig. 7C) when compared with images of the identical area at $T = 1.3$ K (Fig. 7A). Our time resolution was limited to ~ 1 ms so that we were imaging the time-averaged magnetization. Thus, the loss of spectral intensity in the Q_A pockets was most likely caused by increased fluctuations of the magnetization states within the Q_A pockets, which were thermally activated at this temperature. Thus, the fluctuations of the magnetization led to an overall reduction of the measured time-averaged spin polarization. Nevertheless, we observed similar slow aging behavior after application of magnetic field at $T = 4.2$ K for the other Q-pockets as seen at lower temperatures for exactly the same sample area (Fig. 7F). This observation illustrated that there were at least two different and Q-dependent time scales present at sufficient temperature.

Although it is unclear how to describe the Q-dependent dynamical behavior of the magnetization, the presence of multiple relaxation times is an indication of glassy behavior in a mean-field picture, analogous to the aging effects seen in metallic alloys (3, 10). There may be some ruggedness to the energy barriers, symbolic of a multiwell landscape, resulting from strong local order in this system. In struc-

tural glasses, local order leads to a more complex picture of the dynamics, referred to as dynamic heterogeneity (64). In this picture, there can be different spatial regions exhibiting substantially different dynamics related to the local correlations. The existence of different and Q-dependent time scales may be evidence for dynamic heterogeneity in Nd. Therefore, it is not entirely clear how many various time scales can be seen in this system and at what temperatures compared with conventional spin glasses, which have more flat energy landscapes. Combined with the observations that lower defect densities led to a smearing of the Q states, these results exemplify the complex aging behavior related to the multiwell energy landscape illustrated by theory in Fig. 3. The application of higher temperature led to more well-defined Q states (e.g., for $T = 7$ K); by comparison, the same area imaged at $T = 40$ mK showed glassy distributions similar to the previously shown data taken at $T = 1.3$ K (see supplementary text, section S11) (50). This result rules out transitions into various long-range ordered multi- Q states (21, 25) with decreasing temperature and instead exemplifies the emergence of a spin-Q glass state.

Discussion

We have demonstrated that the magnetic ground state of elemental Nd is a spin-Q glass. We have observed aging, i.e., glassiness, in an elemental

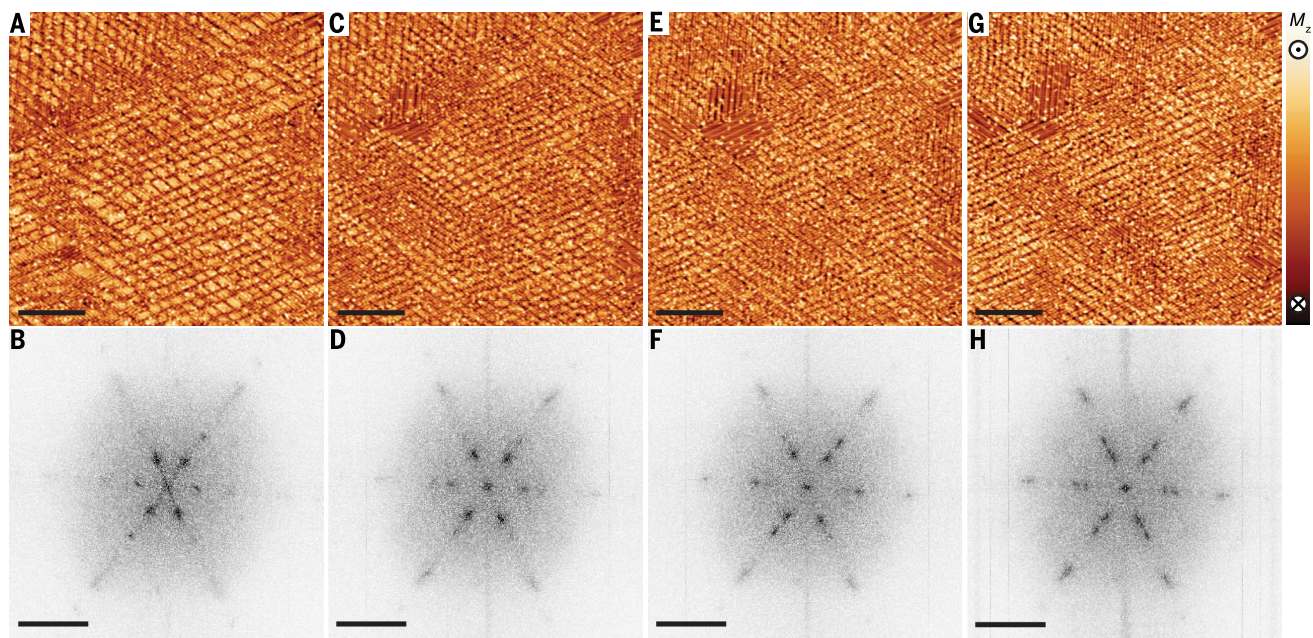


Fig. 6. Aging and glassy behavior of the spin-Q glass state. (A and B) Magnetization and the corresponding Q-space images of the surface at $T = 1.3$ K in its pristine state at $B = 0$ T. The same measurements were performed at the exact same area in $B = 0$ T after subsequent magnetic field sweeps and intermittent probing, leading to the magnetization and Q-space images after $B_{z1} = +4$ T, $\tau_1 = 10^5$ s

(C and D), after $B_{z2} = -4$ T, $\tau_2 = 10^5$ s (E and F), and after $B_{z3} = +7$ T, $\tau_3 = 10^5$ s (G and H). Typical sweep rates were on the order of 225 mT/min. The sequence shows that the system never reverts to the initial zero-field cooled state (scale bar, 30 nm, $I_t = 200$ pA for magnetization images; scale bar, 4 nm^{-1} for Q-space images, inverted grayscale). The surface defect concentration is 0.01 ML.

magnetic solid with minimal amounts of chemical or extrinsic disorder, which is direct experimental confirmation of self-induced glassiness (15). Moreover, the coexistence of short-range order exhibited by multi-Q pockets along with aging behavior in a material without extrinsic disorder cannot be captured by traditional mean-field theoretical descriptions of spin glasses (1, 2). The self-induced glassy behavior observed here is a departure from the traditional magnetic alloys, where disorder drives the glassy dynamics and, likewise, many intermediate time scales can be observed. The energy landscape of Nd may have some amount of ruggedness resulting from the strong local Q order (17), providing a new material system with which to study dynamic heterogeneity in a spin glass material (64). With the advent of techniques that can probe picosecond dynamics with STM (65), it may be interesting to develop magnetically sensitive, time-resolved methods that can resolve the picosecond dynamics in Nd, which can then be compared directly with ASD simulations. It remains to be understood what is particularly special about the interplay between crystal structure and electronic properties in Nd that leads to this exotic behavior, necessitating a deeper theoretical understanding of the role of electron correlation effects, the possible influence of the surface, as well as the interplay between spin and orbital degrees of freedom.

The conclusions drawn here resolve the long-standing debate about the magnetic state of

Nd in that the various reported multi-Q transitions as a function of temperature can be understood within a picture of multiple minima in Q-space with different depths that can be thermally activated. Likewise, the establishment of spin-Q glass order raises the question of the dynamical behavior of spins through the variation in local relaxation times and may provide a material platform with which to explore exotic topological quasiparticles similar to fractons (66–68). The example here expands our views on magnetic states of matter and invites numerous further experimental and theoretical investigations to understand the emergence of aging behavior in magnetic systems.

Materials and methods

The experimental studies were performed in two home-built, ultrahigh vacuum systems that allow for cleaning the W(110) single-crystal substrates, molecular beam epitaxy, and annealing of Nd (see supplementary text, section S1), as well as transfer into the cryogenic STM, all in situ. The first system operates at a base temperature of 1.2 K with magnetic fields up to 9 T perpendicular to the sample. The second system operates at a base temperature of 30 mK with magnetic fields up to 9 T perpendicular and 4 T parallel to the sample (69). The SP-STM measurements were performed using an antiferromagnetic Cr tip with out-of-plane spin polarization (see supplementary text, section S2). Apart from a global plane subtraction, all STM topography images

are unprocessed raw data. Magnetization images were produced by subtraction of the majority and the minority SP-STM images, and corresponding Q-space images were produced by computing the FFT using MATLAB. For the latter, we again did not apply any post-processing steps (see supplementary text, section S2).

The theoretical studies included ab initio calculations of bulk Nd as well as cubic- and hexagonal-terminated slabs with up to 13 layers. The electronic structure was calculated within density functional theory using the RSPt software (70). We used the local spin-density approximation for the exchange and correlation functional. In these calculations, we made use of the standard model of the lanthanides and treated the 4f electrons as localized, unhybridized particles with a magnetic moment according to Russell–Saunders coupling. The dispersive states were expanded by means of 6s, 6p, and 5d orbitals in a multiple-basis fashion, as described in (70). In addition, we included pseudocore 5s and 5p states to hybridize with the rest of the valence states. No approximation was made concerning the shape of the charge density or potential in a so-called full-potential description. The simulations were done using 32,768 and 1152 k-points of the full BZ for bulk dhcp and slabs, respectively (see supplementary text, section S4). The Monte Carlo and ASD calculations were performed using The Uppsala Atomic Spin Dynamics (UppASD) software (71, 72) (see supplementary text, section S5).

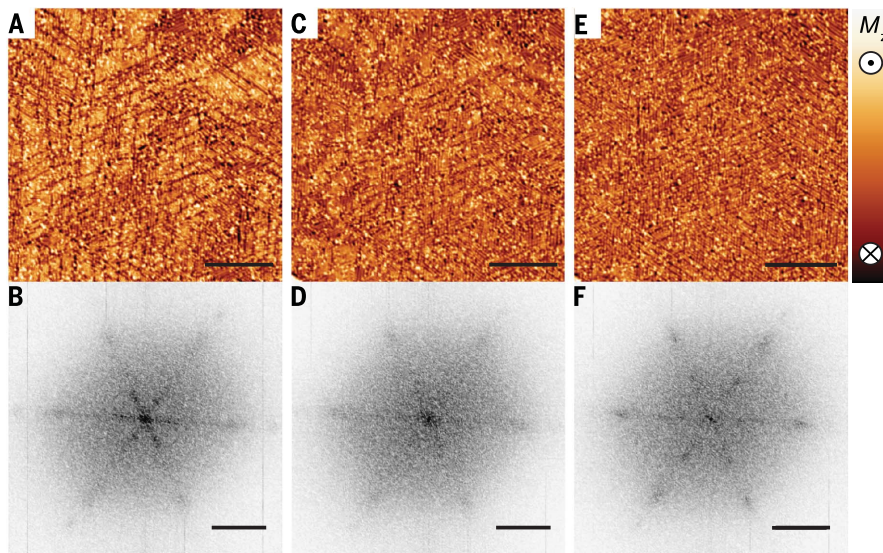


Fig. 7. Temperature dependence of the spin-Q glass state. Magnetization and the corresponding Q-space images of the exact same area at $T = 1.3$ K (A and B) and at $T = 4.2$ K (C and D) in $B = 0$ T. Warming up the surface from 1.3 to 4.2 K results in depopulation of the Q_A pocket. (E and F) The same measurements were performed after a +4 T magnetic field sweep in the same area at $T = 4.2$ K in $B = 0$ T, illustrating similar out-of-plane magnetic field aging behavior (scale bar, 30 nm, $I_t = 200$ pA for magnetization images; scale bar, 3 nm^{-1} for Q-space images, inverted grayscale). The surface defect concentration is 0.01 ML.

REFERENCES AND NOTES

- S. F. Edwards, P. W. Anderson, Theory of spin glasses. *J. Phys. F Met. Phys.* **5**, 965–974 (1975). doi: [10.1088/0305-4608/5/5/017](#)
- S. F. Edwards, P. W. Anderson, Theory of spin glasses. II. *J. Phys. F Met. Phys.* **6**, 1927–1937 (1976). doi: [10.1088/0305-4608/6/10/022](#)
- K. Binder, A. P. Young, Spin glasses: Experimental facts, theoretical concepts, and open questions. *Rev. Mod. Phys.* **58**, 801–976 (1986). doi: [10.1103/RevModPhys.58.801](#)
- M. Mézard, G. Parisi, M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, 1987).
- K. H. Fischer, J. A. Hertz, *Spin Glasses* (Cambridge Univ. Press, 1993).
- N. Katayama *et al.*, Investigation of the spin-glass regime between the antiferromagnetic and superconducting phases in $\text{Fe}_{1-x}\text{Se}_x\text{Te}_{1-x}$. *J. Phys. Soc. Jpn.* **79**, 113702 (2010). doi: [10.1143/JPSJ.79.113702](#)
- J. Hertz, A. Krough, R. G. Palmer, *Introduction to the Theory of Neural Computation* (CRC, 1991).
- Y. Zhou, K. Kanoda, T.-K. Ng, Quantum spin liquid states. *Rev. Mod. Phys.* **89**, 025003 (2017). doi: [10.1103/RevModPhys.89.025003](#)
- S. T. Bramwell, M. J. P. Gingras, Spin ice state in frustrated magnetic pyrochlore materials. *Science* **294**, 1495–1501 (2001). doi: [10.1126/science.1064761](#); pmid: [11711667](#)
- B. Skubic *et al.*, Atomistic spin dynamics of the Cu-Mn spin-glass alloy. *Phys. Rev. B* **79**, 024411 (2009). doi: [10.1103/PhysRevB.79.024411](#)
- R. Rammal, G. Toulouse, M. A. Virasoro, Ultrametricity for physicists. *Rev. Mod. Phys.* **58**, 765–788 (1986). doi: [10.1103/RevModPhys.58.765](#)
- M. Mézard, G. Parisi, A tentative replica study of the glass transition. *J. Phys. A Math. Theor.* **29**, 6515–6524 (1996).
- J. Schmalian, P. G. Wolynes, Stripe glasses: Self-generated randomness in a uniformly frustrated system. *Phys. Rev. Lett.* **85**, 836–839 (2000). doi: [10.1103/PhysRevLett.85.836](#); pmid: [10991411](#)
- H. Westfahl, J. Schmalian, P. G. Wolynes, Self-generated randomness, defect wandering, and viscous flow in stripe glasses. *Phys. Rev. B* **64**, 174203 (2001). doi: [10.1103/PhysRevB.64.174203](#)
- A. Principi, M. I. Katsnelson, Self-Induced Glassiness and Pattern Formation in Spin Systems Subject to Long-Range Interactions. *Phys. Rev. Lett.* **117**, 137201 (2016). doi: [10.1103/PhysRevLett.117.137201](#); pmid: [27715081](#)
- A. Principi, M. I. Katsnelson, Stripe glasses in ferromagnetic thin films. *Phys. Rev. B* **93**, 054410 (2016). doi: [10.1103/PhysRevB.93.054410](#)
- A. Kolmus, M. I. Katsnelson, A. A. Khajetoorians, H. J. Kappen, Atom-by-atom construction of attractors in a tunable finite size spin array. *New J. Phys.* **22**, 023038 (2020). doi: [10.1088/1367-2630/ab6f91](#)
- R. M. Moon, J. W. Cable, W. C. Koehler, Magnetic structure of neodymium. *J. Appl. Phys.* **35**, 1041–1042 (1964). doi: [10.1063/1.1713369](#)
- P. Bak, B. Lebech, Triple-Q-modulated magnetic-structure and critical behavior of neodymium. *Phys. Rev. Lett.* **40**, 800–803 (1978). doi: [10.1103/PhysRevLett.40.800](#)
- B. Lebech, J. Alsnielsen, K. A. McEwen, X-ray and neutron-scattering study of the magnetic-structure of neodymium metal. *Phys. Rev. Lett.* **43**, 65–67 (1979). doi: [10.1103/PhysRevLett.43.65](#)
- E. M. Forgan, E. P. Gibbons, K. A. McEwen, D. Fort, Observation of a quadruple-q magnetic structure in neodymium. *Phys. Rev. Lett.* **62**, 470–473 (1989). doi: [10.1103/PhysRevLett.62.470](#); pmid: [10040241](#)
- K. A. McEwen, S. W. Zochowski, Magnetic phase-transitions in neodymium. *J. Magn. Magn. Mater.* **90–91**, 94–98 (1990). doi: [10.1016/S0304-8853\(90\)80035-6](#)
- S. W. Zochowski, K. A. McEwen, E. Fawcett, Magnetic phase-diagrams of neodymium. *J. Phys. Condens. Matter* **3**, 8079–8094 (1991). doi: [10.1088/0953-8984/3/41/007](#)
- E. M. Forgan *et al.*, Field effects on the antiferromagnetic ordering of neodymium. *J. Magn. Magn. Mater.* **104–107**, 911–912 (1992). doi: [10.1016/0304-8853\(92\)90419-0](#)
- B. Lebech, J. Wolny, R. M. Moon, Magnetic phase-transitions in double hexagonal close-packed neodymium metal commensurate in 2 dimensions. *J. Phys. Condens. Matter* **6**, 5201–5222 (1994). doi: [10.1088/0953-8984/6/27/029](#)
- P. A. Lindgard, T. Chatterji, K. Prokes, V. Sikolenko, J. U. Hoffmann, Magnetic diffuse scattering from Nd above T-N and deduced exchange interaction parameters. *J. Phys. Condens. Matter* **19**, 286201 (2007). doi: [10.1088/0953-8984/19/28/286201](#)
- D. Wegner, A. Bauer, A. Rehbein, G. Kaindl, Exchange splittings of lanthanide (0001)-surface states and their dependences on short-range magnetic order. *Jpn. J. Appl. Phys.* **45** (3B), 1941–1945 (2006). doi: [10.1143/JJAP.45.1941](#)
- S. Heinze *et al.*, Spontaneous atomic-scale magnetic skyrmion lattice in two dimensions. *Nat. Phys.* **7**, 713–718 (2011). doi: [10.1038/nphys2045](#)
- D. Weller, S. F. Alvarado, Preparation of Remanently Ferromagnetic Gd(0001). *J. Appl. Phys.* **59**, 2908–2913 (1986). doi: [10.1063/1.336951](#)
- S. C. Wu *et al.*, Surface state on clean Tb(0001). *Phys. Rev. B* **44**, 13720–13724 (1991). doi: [10.1103/PhysRevB.44.13720](#); pmid: [9999578](#)
- G. H. Fecher *et al.*, Dichroism in angular resolved VUV-photoemission from the (0001) surfaces of thin Gd and Nd films epitaxially grown on W(110). *Eur. Phys. J. B* **11**, 161–175 (1999). doi: [10.1007/s100510050926](#)
- E. Weschke, G. Kaindl, 4f- and surface-electronic structure of lanthanide metals. *J. Electron Spectrosc. Relat. Phenom.* **75**, 233–244 (1995). doi: [10.1016/0368-2048\(95\)02540-5](#)
- K. M. Döbrich *et al.*, Fermi-surface topology and helical antiferromagnetism in heavy lanthanide metals. *Phys. Rev. Lett.* **104**, 246401 (2010). doi: [10.1103/PhysRevLett.104.246401](#); pmid: [20867317](#)
- C. Dufour *et al.*, Epitaxial growth of dhcp samarium: Single crystal films and Sm/Nd superlattices. *J. Cryst. Growth* **234**, 447–453 (2002). doi: [10.1016/S0022-0248\(01\)01719-5](#)
- O. Waller *et al.*, Lattice dynamics of neodymium: Influence of 4 f electron correlations. *Phys. Rev. B* **94**, 014303 (2016). doi: [10.1103/PhysRevB.94.014303](#)
- E. Weschke, G. Kaindl, Magnetic exchange splitting in lanthanide metals. *J. Phys. Condens. Matter* **13**, 11133–11148 (2001). doi: [10.1088/0953-8984/13/49/303](#)
- J. Quinn, Y. S. Li, F. Jona, D. Fort, Atomic structure of a Gd(0001) surface. *Phys. Rev. B* **46**, 9694–9699 (1992). doi: [10.1103/PhysRevB.46.9694](#); pmid: [10002781](#)
- J. P. Goff, C. Bryn-Jacobsen, D. F. McMorro, R. C. Ward, M. R. Wells, Formation of uniform magnetic structures and epitaxial hydride phases in Nd/Pr superlattices. *Phys. Rev. B* **55**, 12537–12545 (1997). doi: [10.1103/PhysRevB.55.12537](#)
- E. D. Tober, R. X. Ynzunza, C. Westphal, C. S. Fadley, Relationship between morphology and magnetic behavior for Gd thin films on W(110). *Phys. Rev. B* **53**, 5444–5448 (1996). doi: [10.1103/PhysRevB.53.5444](#); pmid: [9984151](#)
- S. Jo, Y. Gotoh, Growth modes of Nd deposited on Mo(110) surface. *Jpn. J. Appl. Phys.* **38** (Part 1, No. 12A), 6878–6881 (1999). doi: [10.1143/JJAP.38.6878](#)
- H. Li *et al.*, Structural and electronic properties of ultrathin films of Gd, Tb, Dy, Ho, and Er. *Phys. Rev. B* **45**, 3853–3856 (1992). doi: [10.1103/PhysRevB.45.3853](#); pmid: [10001981](#)
- M. Farle, K. Baberschke, U. Stetter, A. Aspelmeier, F. Gerhardter, Thickness-dependent Curie temperature of Gd(0001)/W(110) and its dependence on the growth conditions. *Phys. Rev. B* **47**, 11571–11574 (1993). doi: [10.1103/PhysRevB.47.11571](#); pmid: [10005313](#)
- C. S. Arnold, D. P. Pappas, Gd(0001): A semi-infinite three-dimensional heisenberg ferromagnet with ordinary surface transition. *Phys. Rev. Lett.* **85**, 5202–5205 (2000). doi: [10.1103/PhysRevLett.85.5202](#); pmid: [11102221](#)
- C. Schussler-Langeheine *et al.*, Magnetic splitting of valence states in ferromagnetic and antiferromagnetic lanthanide metals. *Phys. Rev. Lett.* **84**, 5624–5627 (2000). doi: [10.1103/PhysRevLett.84.5624](#); pmid: [10991010](#)
- E. Weschke *et al.*, Finite-size effect on magnetic ordering temperatures in long-period antiferromagnets: Holmium thin films. *Phys. Rev. Lett.* **93**, 157204 (2004). doi: [10.1103/PhysRevLett.93.157204](#); pmid: [15524935](#)
- B. A. Everitt *et al.*, Helimagnetic structures in epitaxial Nd/Y superlattices and alloys. *Phys. Rev. B* **56**, 5452–5460 (1997). doi: [10.1103/PhysRevB.56.5452](#)
- J. P. Goff *et al.*, Magnetism and superconductivity in neodymium/lanthanum superlattices. *Physica B Condens. Matter* **241–243**, 714–716 (1997). doi: [10.1016/S0921-4526\(97\)00701-1](#)
- A. Aspelmeier, F. Gerhardter, K. Baberschke, Magnetism and structure of ultrathin Gd films. *J. Magn. Magn. Mater.* **132**, 22–30 (1994). doi: [10.1016/0304-8853\(94\)90296-8](#)
- M. Bode, M. Getzlaff, S. Heinze, R. Pascal, R. Wiesendanger, Magnetic exchange splitting of the Gd(0001) surface state studied by variable-temperature scanning tunneling spectroscopy. *Appl. Phys. A Mater. Sci. Process.* **66**, S121–S124 (1998). doi: [10.1007/s003390051113](#)
- See the supplementary materials.
- S. Krause *et al.*, Consequences of line defects on the magnetic structure of high anisotropy films: Pinning centers on Dy/W(110). *Europhys. Lett.* **76**, 637–643 (2006). doi: [10.1209/epl/2006-10307-2](#)

52. M. Getzlaff, M. Bode, R. Pascal, R. Wiesendanger, Adsorbates on Gd(0001): A combined scanning tunneling microscopy and photoemission study. *Phys. Rev. B* **59**, 8195–8208 (1999). doi: [10.1103/PhysRevB.59.8195](https://doi.org/10.1103/PhysRevB.59.8195)
53. D. Wegner, A. Bauer, G. Kaindl, Effect of impurities on Tamm-like lanthanide-metal surface states. *Phys. Rev. B* **76**, 113410 (2007). doi: [10.1103/PhysRevB.76.113410](https://doi.org/10.1103/PhysRevB.76.113410)
54. D. Wegner, A. Bauer, G. Kaindl, Magnon-broadening of exchange-split surface states on lanthanide metals. *Phys. Rev. B* **73**, 165415 (2006). doi: [10.1103/PhysRevB.73.165415](https://doi.org/10.1103/PhysRevB.73.165415)
55. T. Hanaguri *et al.*, A 'checkerboard' electronic crystal state in lightly hole-doped $\text{Ca}_2\text{-xNaxCuO}_2\text{Cl}_2$. *Nature* **430**, 1001–1005 (2004). doi: [10.1038/nature02861](https://doi.org/10.1038/nature02861); pmid: [15329714](https://pubmed.ncbi.nlm.nih.gov/15329714/)
56. S. Mühlbauer *et al.*, Skyrmion lattice in a chiral magnet. *Science* **323**, 915–919 (2009). doi: [10.1126/science.1166767](https://doi.org/10.1126/science.1166767); pmid: [19213914](https://pubmed.ncbi.nlm.nih.gov/19213914/)
57. X. Z. Yu *et al.*, Real-space observation of a two-dimensional skyrmion crystal. *Nature* **465**, 901–904 (2010). doi: [10.1038/nature09124](https://doi.org/10.1038/nature09124); pmid: [20559382](https://pubmed.ncbi.nlm.nih.gov/20559382/)
58. K. von Bergmann, M. Menzel, A. Kubetzka, R. Wiesendanger, Influence of the local atom configuration on a hexagonal skyrmion lattice. *Nano Lett.* **15**, 3280–3285 (2015). doi: [10.1021/acs.nanolett.5b00506](https://doi.org/10.1021/acs.nanolett.5b00506); pmid: [25859818](https://pubmed.ncbi.nlm.nih.gov/25859818/)
59. W. C. Koehler, Magnetic properties of rare-earth metals and alloys. *J. Appl. Phys.* **36**, 1078–1087 (1965). doi: [10.1063/1.1714108](https://doi.org/10.1063/1.1714108)
60. I. L. M. Locht *et al.*, Standard model of the rare earths analyzed from the Hubbard I approximation. *Phys. Rev. B* **94**, 085137 (2016). doi: [10.1103/PhysRevB.94.085137](https://doi.org/10.1103/PhysRevB.94.085137)
61. Y. O. Kvashnin *et al.*, Exchange parameters of strongly correlated materials: Extraction from spin-polarized density functional theory plus dynamical mean-field theory. *Phys. Rev. B* **91**, 125133 (2015). doi: [10.1103/PhysRevB.91.125133](https://doi.org/10.1103/PhysRevB.91.125133)
62. P. Löptien, L. Zhou, A. A. Khajetoorians, J. Wiebe, R. Wiesendanger, Superconductivity of lanthanum revisited: Enhanced critical temperature in the clean limit. *J. Phys. Condens. Matter* **26**, 425703 (2014). doi: [10.1088/0953-8984/26/42/425703](https://doi.org/10.1088/0953-8984/26/42/425703); pmid: [25272968](https://pubmed.ncbi.nlm.nih.gov/25272968/)
63. N. Romming *et al.*, Writing and deleting single magnetic skyrmions. *Science* **341**, 636–639 (2013). doi: [10.1126/science.1240573](https://doi.org/10.1126/science.1240573); pmid: [23929977](https://pubmed.ncbi.nlm.nih.gov/23929977/)
64. M. D. Ediger, Spatially heterogeneous dynamics in supercooled liquids. *Annu. Rev. Phys. Chem.* **51**, 99–128 (2000). doi: [10.1146/annurev.physchem.51.1.99](https://doi.org/10.1146/annurev.physchem.51.1.99); pmid: [11031277](https://pubmed.ncbi.nlm.nih.gov/11031277/)
65. T. L. Cocker, D. Peller, P. Yu, J. Repp, R. Huber, Tracking the ultrafast motion of a single molecule by femtosecond orbital imaging. *Nature* **539**, 263–267 (2016). doi: [10.1038/nature19816](https://doi.org/10.1038/nature19816); pmid: [27830788](https://pubmed.ncbi.nlm.nih.gov/27830788/)
66. H. Rieger, A. P. Young, in *Complex Behaviour of Glassy Systems*, M. Rubí, C. Pérez-Vicente, Eds. (Springer, 1997).
67. C. Chamon, Quantum glassiness in strongly correlated clean systems: An example of topological overprotection. *Phys. Rev. Lett.* **94**, 040402 (2005). doi: [10.1103/PhysRevLett.94.040402](https://doi.org/10.1103/PhysRevLett.94.040402); pmid: [15783534](https://pubmed.ncbi.nlm.nih.gov/15783534/)
68. R. M. Nandkishore, M. Hermele, Fractons. *Annu. Rev. Condens. Matter Phys.* **10**, 295–313 (2019). doi: [10.1146/annurev-conmatphys-031218-013604](https://doi.org/10.1146/annurev-conmatphys-031218-013604)
69. H. von Allwörden *et al.*, Design and performance of an ultra-high vacuum spin-polarized scanning tunneling microscope operating at 30 mK and in a vector magnetic field. *Rev. Sci. Instrum.* **89**, 033902 (2018). doi: [10.1063/1.5020045](https://doi.org/10.1063/1.5020045); pmid: [29604794](https://pubmed.ncbi.nlm.nih.gov/29604794/)
70. J. M. Wills *et al.*, *Full-Potential Electronic Structure Method* (Springer, 2010).
71. Uppsala Atomic Spin Dynamics (UppASD) software; <https://github.com/UppASD/UppASD>
72. O. Eriksson, A. Bergman, L. Bergqvist, J. Hellsvik, *Atomistic Spin Dynamics: Foundations and Applications* (Oxford Univ. Press, 2017).

ACKNOWLEDGMENTS

We thank P. Thunström for assistance with the spin-polarized core calculations. **Funding:** This work was supported by the Swedish National Infrastructure for Computing (SNIC), NWO, and VIDI project no. 680-47-534, "Manipulating the interplay between superconductivity and chiral magnetism at the single-atom level." This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (SPINAPSE; grant no. 818399). We also acknowledge support from the Swedish Research Council (VR), the Knut and Alice Wallenberg Foundation (KAW), the Foundation for Strategic Research (SSF), Energimyndigheten, eSENCE, and StandUP. **Author contributions:** U.K., A.E., and M.S. conducted the experiments. U.K., A.E., M.S., N.H. D.W., and A.A.K. analyzed the experimental data. A.A.K. and D.W. designed the experiments. D.I. performed ab initio calculations. A.B. performed the spin-dynamics simulations. M.I.K., L.N., and O.E. provided additional theoretical support on all calculations. All authors contributed to the writing of the manuscript. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** All data needed to evaluate the conclusions of this study are available in the main text or the supplementary materials. Any additional data pertaining to this study can be made available, pending approval, by contacting the corresponding author.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6494/eaay6757/suppl/DC1
Supplementary Text
Figs. S1 to S16
Tables S1 and S2
References (73–100)

9 July 2019; resubmitted 14 January 2020
Accepted 10 April 2020
[10.1126/science.aay6757](https://doi.org/10.1126/science.aay6757)

RESEARCH ARTICLE

INVASIVE SPECIES

Biotic interactions drive ecosystem responses to exotic plant invaders

L. P. Waller^{1*}, W. J. Allen², B. I. P. Barratt^{3,4}, L. M. Condon¹, F. M. França^{5,6}, J. E. Hunt⁷, N. Koele⁷, K. H. Orwin⁷, G. S. Steel², J. M. Tylianakis², S. A. Wakelin⁸, I. A. Dickie^{2*}

Ecosystem process rates typically increase after plant invasion, but the extent to which this is driven by (i) changes in productivity, (ii) exotic species' traits, or (iii) novel (non-coevolved) biotic interactions has never been quantified. We created communities varying in exotic plant dominance, plant traits, soil biota, and invertebrate herbivores and measured indicators of carbon cycling. Interactions with soil biota and herbivores were the strongest drivers of exotic plant effects, particularly on measures of soil carbon turnover. Moreover, plant traits related to growth and nutrient acquisition explained differences in the ways that exotic plants interacted with novel biota compared with natives. We conclude that novel biological interactions with exotic species are a more important driver of ecosystem transformation than was previously recognized.

Introductions of exotic plants are transforming Earth's ecosystems and altering the way in which they cycle carbon (C) (1–7). Successful exotic species can modify C-cycling through several potential inter-related mechanisms, most notably by increases to community productivity driving higher amounts of biomass returned to soil (2) and different traits of exotics compared with native species driving a higher proportion of high-quality biomass to soil (1, 5, 8). Much less attention has focused on how altered biotic interactions in an exotic species' new range may affect C-cycling. It is well known that organisms above and belowground such as plants, soil biota, and invertebrate herbivores mediate ecosystem processes related to C-cycling and that exotic plants often differ in their interactions with new biota where there is no shared co-evolutionary history (i.e., novel) (5–7, 9). Thus, a third potential mechanism is that differences in the ways that exotics interact with novel organisms in recipient communities explain changes to C-cycling. However, the relative importance of, and interactions among, these three mechanisms are unknown; most studies have focused on systems invaded by species with high productivity, making it impossible to separate the impacts of community biomass increases from other mechanisms.

Both predictive understanding and applied management of C stocks require these three nonmutually exclusive hypothesized mechanisms to be disentangled. Here, we resolve this gap in knowledge with a multifactor, multi-response experiment aimed at partitioning the drivers of exotic impacts on ecosystem functions related to C-cycling.

We established 160 experimental ecosystems (mesocosms); manipulated interactions among plants, soil biota, and invertebrate herbivores

in a fully factorial design; and measured how the traits and biomass of exotic species and biotic interactions affected ecosystem properties and processes (10) (Fig. 1A). Each mesocosm was grown in a 125-L pot (575-mm diameter; Fig. 1B) and comprised one of 20 individual, eight-species plant communities (table S1) varying orthogonally in the proportion of exotic and woody shrub and/or tree species (0 to 100% and 0 to 63%, respectively). These plants were taken from a pool of 20 exotic and 19 native and/or endemic New Zealand plant species (table S2). Soil biota were manipulated using a modified plant–soil feedback approach (11). Each plant species was grown in monoculture in 10-L pots containing field-collected soil for 9 to 10 months (for soil collection details, see table S3), allowing the conditioning of typical associated soil biota for each of the plant species (Fig. 1C). We created “home” soils by taking the conditioned soil from each of the eight representative species in a mesocosm and mixing it together to create a single inoculum. Each home soil mixture was also used as an “away” soil in a different mesocosm that did not contain any of the representative plants in that inoculum (table S4). These soils were intended to increase the relative biomass in inocula of specialized and preferred interaction partners of the resident (or nonresident) plant species, including but not limited to bacteria, fungi, oomycetes, and nematodes. Invertebrate insect herbivore populations were added into half of the mesocosms with home soils and

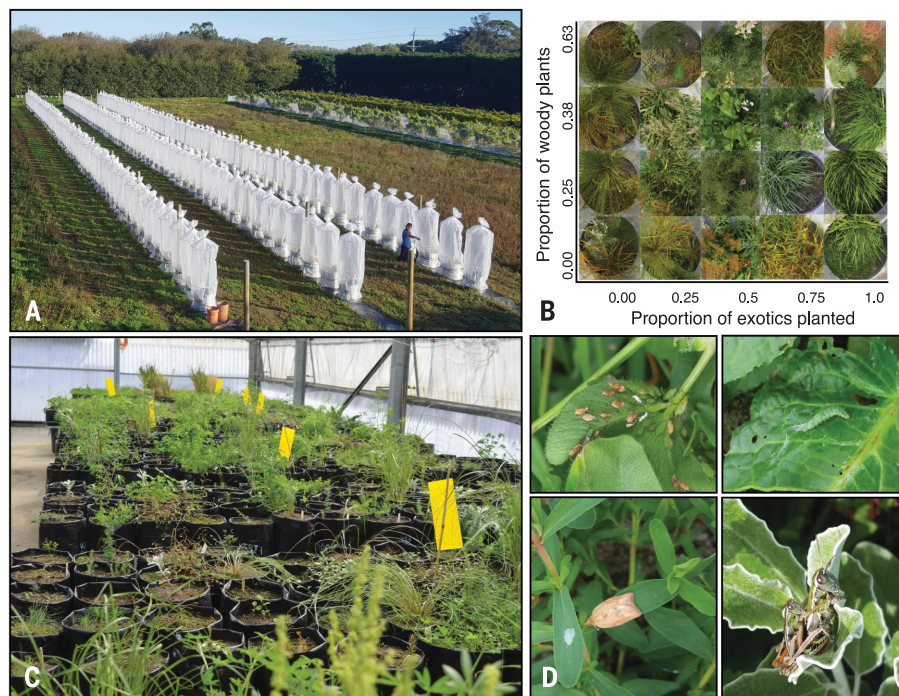


Fig. 1. Details of the experiment. (A) Aerial view and (B) orthogonal design of the 20 communities. (C) Plants grown in the initial soil-conditioning stage. (D) Four of the 20 herbivores established in the mesocosms.

¹Bio-Protection Research Centre, Lincoln University, Lincoln 7647, New Zealand. ²Bio-Protection Research Centre, School of Biological Sciences, University of Canterbury, Christchurch 8041, New Zealand. ³AgResearch, Invermay Research Centre, Mosgiel 9092, New Zealand. ⁴Department of Botany, University of Otago, Dunedin 9016, New Zealand. ⁵Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK. ⁶Embrapa Amazônia Oriental, Belém 66095-100, Brazil. ⁷Manaaki Whenua Landcare Research, Lincoln 7640, New Zealand. ⁸Scion, Christchurch 8011, New Zealand.

*Corresponding author. Email: laurenwallergmail.com (L.P.W.); Ian.Dickie@canterbury.ac.nz (I.A.D.)

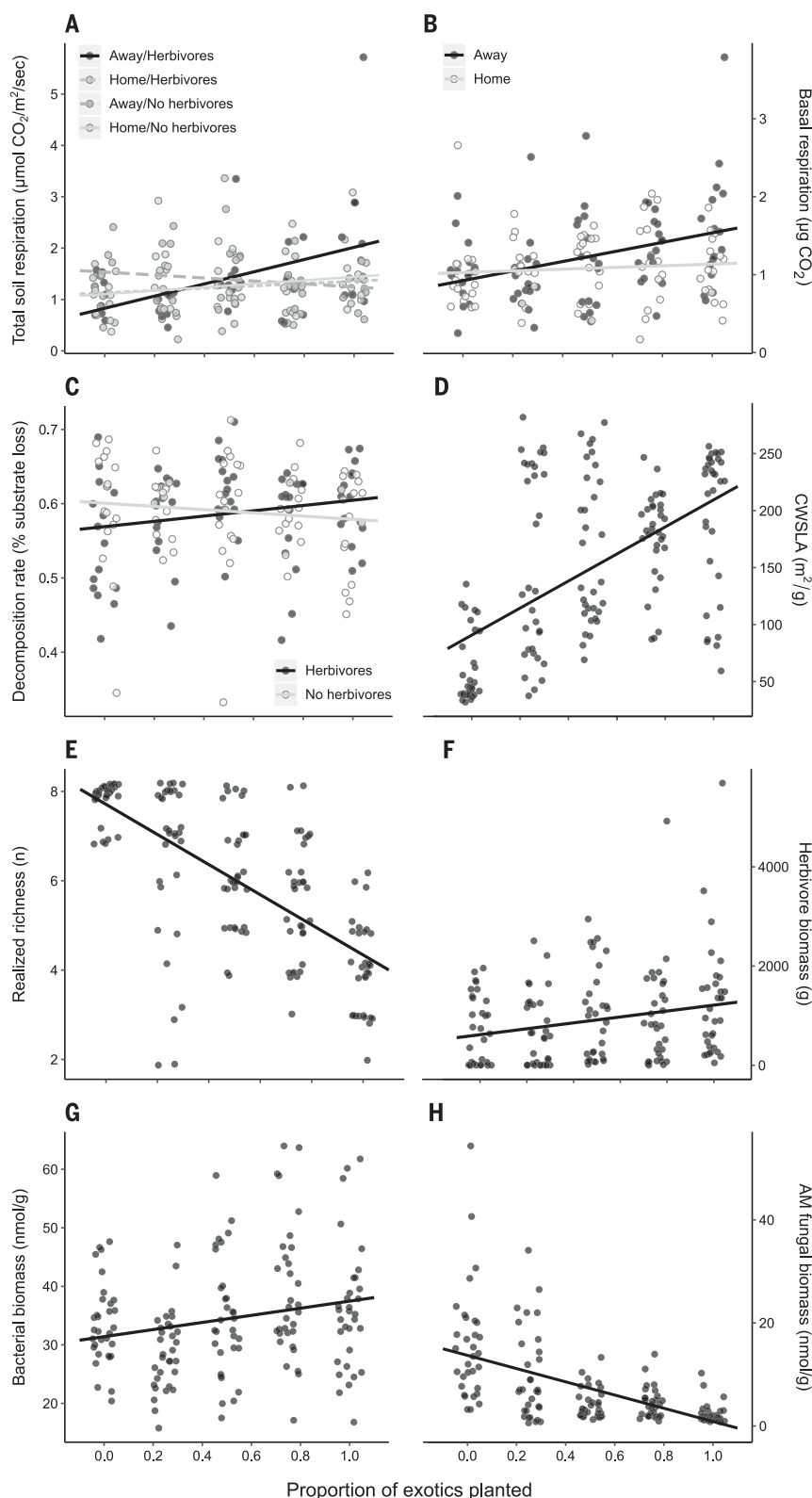


Fig. 2. Relationship between exotic plants and ecosystem properties. (A) Total soil respiration. (B) Basal respiration. (C) Decomposition rate. (D) Community-weighted SLA. (E) Realized plant richness. (F) Herbivore biomass. (G) Bacterial biomass and (H) AMF biomass as a function of proportion of exotics planted and our imposed biotic treatments. Only significant predictors are shown; nonsignificant predictors were removed from the full model based on Akaike information criterion (AIC) selection. Regression lines represent the output of the best-fitting mixed model.

half with away soils. Thirteen invertebrate herbivore species (hereafter simply referred to as “herbivores”) introduced into the mesocosms successfully established, along with seven self-colonizing species, totaling 20 species in all (table S5). All mesocosms were sealed with mesh cages (15% shade factor) designed to retain added herbivores and exclude others from entering (Fig. 1, A and D).

In each mesocosm, we measured distinct ecosystem properties and processes that are relevant to C-cycling dynamics (12): above- and belowground biomass; total in situ soil respiration (R_T); basal respiration (R_B , a measure of soil microbial activity); microbial biomass (R_{St} , measured as substrate-induced respiration); decomposition (rate of standardized substrate mass loss); soil organic matter; nitrogen (N) availability; herbivore biomass; arbuscular mycorrhizal fungi (AMF) biomass and other fungal biomass [using neutral lipid fatty acid (NLFA) and phospholipid fatty acid (PLFA) biomarkers, respectively]; and bacterial biomass (using PLFA). We measured plant traits representative of fundamental plant life history strategies (13): specific leaf area (SLA), specific root length (SRL), fine-root production, the rate of above- and belowground growth [net primary productivity (NPP)], and the rate of increase in plant height between two sampling periods. We also collected leaf N content (N_{mass}) data from an external database where available (table S2). SLA is an easily measured trait that correlates positively with relative growth rate (13) and nutrient use efficiency (14) and can directly influence soil C-cycling (15). Root traits such as SRL can directly affect interactions belowground with soil biota that drive changes to ecosystem processes (16). SLA and SRL were measured for each species separately and then scaled to the community level by calculating community-weighted mean trait values using plant biomass from each mesocosm (10). We incorporated more N-fixing exotic species into our communities compared with native species because N-fixers are more common among the New Zealand exotic flora compared with the native flora, and we would expect N-fixing and non-N-fixing exotics to differ in their effects on ecosystem properties (2). Likewise, we might expect differences between herbaceous versus woody species in their impact on functioning (2), so we included a woody gradient orthogonal to proportion exotic.

We used linear mixed-effects models to test our hypothesis that measured ecosystem properties were affected by our treatments, with the proportion of exotic plants in a community, soil and herbivore treatments, interactions among these factors, and plant effect traits (i.e., SLA, weighted by community biomass) as covariates. We used piecewise structural equation modeling (SEM) to explore how exotic species and their traits directly and indirectly

(through modifying herbivore biomass, soil microbes, and soil nutrients) influenced C-cycling. To draw the causal links between exotic plant dominance (defined here as a higher richness of exotic compared with native plants), plant traits, herbivores, and soil microbiota to our ecosystem functions in the SEM, we modeled our dependent variables as a function of the experimental responses measured in the mesocosms. We measured SLA and SRL from plants outside of the mesocosms to capture effect traits (17) that were broadly representative of the life history strategies of our plant species but not the plastic responses to our treatments. Thus, even though we weighted trait values by biomass of each plant in the mesocosms, the community-weighted SLA and

SRL values used in the hypothesized models and the SEM may over- or underestimate the response of these traits. Finally, to partition the effects of N-fixing exotic species versus non-N-fixing exotics from different plant functional groups, we included the realized proportion of woody and herbaceous N-fixing and non-N-fixing exotic plants into the SEM.

Altered biotic interactions are the strongest drivers of ecosystem functions

Biotic interactions (herbivory and home and/or away soil treatments) modified the impact of exotic species on measures of soil C-cycling (i.e., respiration and decomposition). The presence of herbivores and away soil doubled in situ soil respiration rates in exotic-dominated

communities compared with corresponding native communities (R_T ~proportion exotic*soil origin*herbivore, $\chi^2 = 10.41$, $p < 0.001$; Fig. 2A). Reduced abundance of species-specific soil biota (away soil) in exotic-dominated communities increased soil microbial respiration by 1.5 times compared with the same communities grown in soils cultured to contain more specialist biota (R_B ~proportion exotic*soil origin, $\chi^2 = 4.77$, $p < 0.03$; Fig. 2B). Soil C:N ratios were also lower in away soils under exotic plants (C:N~proportion exotic*soil, $\chi^2 = 3.87$, $p < 0.05$). Higher proportions of N-fixing plant biomass also explained increases in basal respiration in exotic-dominated communities (Figs. 3 and 4). Herbivore feeding on exotic plants correlated positively with decomposition rates (decomposition~proportion exotic*herbivore, $\chi^2 = 4.29$, $p < 0.04$; Fig. 2C). These results underpin the need to explicitly include biological interactions in C-cycling models to improve predictions of ecosystem and global C budgets.

Ecosystem effects by exotic plants are mediated by leaf traits, not increases to community biomass

Exotic species had 1.8 times higher mean SLA (SLA~plant provenance, $\chi^2 = 44.64$, $p < 0.001$; fig. S3), grew taller 2.5 times faster (height~plant provenance, $\chi^2 = 49.42$, $p < 0.0001$), had 2 times higher leaf N (N_{mass} ~plant provenance, $\chi^2 = 5.05$, $p < 0.02$), and assimilated twice as much C into tissues compared with native plants (NPP_{shoots} ~plant provenance, $\chi^2 = 5.24$, $p < 0.02$; NPP_{roots} ~plant provenance, $\chi^2 = 3.61$, $p = 0.057$), congruent with traits of exotic plants worldwide (8). Root traits (SRL, average diameter of fine roots and the proportion of fine roots), however, did not differ between native and exotic species (SRL~plant provenance, $\chi^2 = 0.25$, $p = 0.62$; average diameter, $\chi^2 = 0.32$, $p = 0.57$; fine roots, $\chi^2 = 0.05$, $p = 0.82$). Exotic plants also increased mean leaf trait values at the community level, because the community-weighted mean SLA (CWSLA) increased with the proportion of exotics in the community (CWSLA~proportion exotic, $\chi^2 = 12.64$, $p < 0.001$; Fig. 2D). High SLA is indicative of a suite of traits allowing for faster growth (13), potentially explaining the larger biomass and faster growth of exotic compared with native plants.

Exotic invaders can increase community biomass as a function of higher resource acquisition (18) and resource use efficiency (19), stronger competitive ability [(20); but not always, see (21) for exceptions], and ability to reach high abundance (22). In our experiment, exotic plants were indeed larger and more competitive than native plants, as evidenced by double the per capita plant biomass (per capita plant biomass~plant provenance, $\chi^2 = 7.66$, $p < 0.006$; fig. S1) and a drop in total plant richness as the proportion of planted

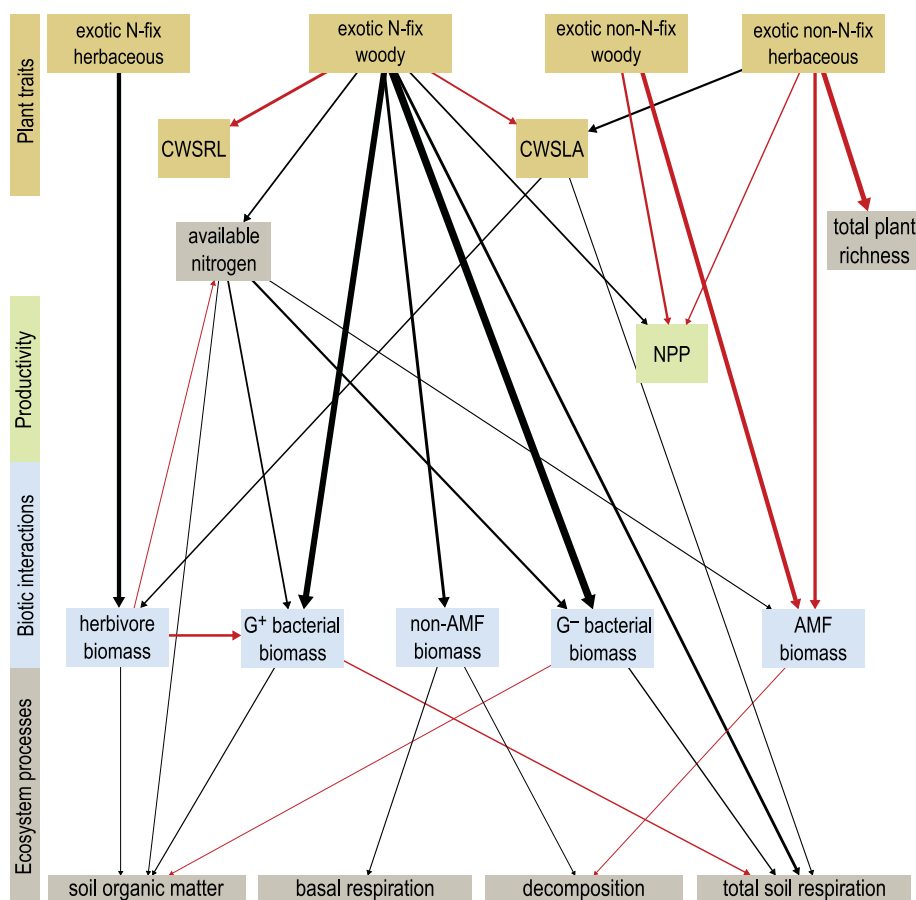


Fig. 3. Piecewise structural equation model showing the inferred direct and indirect effects of exotic plant richness, traits and biotic interactions on indicators of ecosystem function. Plant traits (yellow boxes) are the realized proportional richness of exotic N-fixing and non-N-fixing woody and herbaceous plants, CWSLA, and community-weighted SRL (CWSRL). Productivity (green boxes) is measured as NPP, the rate of plant biomass produced over the course of the experiment; Biotic interactions (blue boxes) include AMF biomass, non-AMF biomass, herbivore biomass, and Gram-positive (G^+) and Gram-negative (G^-) bacterial biomass. Ecosystem processes (gray boxes) are the total plant richness, available N, decomposition, soil organic matter, basal respiration, and total in situ respiration. A priori hypothesized models can be found in table S6. Black arrows indicate significant factors associated with positive effects on response variables and red arrows indicate negative effects. Only significant predictors are shown; nonsignificant predictors were removed from the full model on the basis of AIC selection. Arrow width corresponds to absolute values of relative effect sizes, which can be found, along with standardized coefficients and p values for each path, in table S7.

exotic species increased (plant richness~proportion exotic, $\chi^2 = 47.36$, $p < 0.001$; Fig. 2E), likely through competitive exclusion. Although species richness declined along the exotic gradient, exotic plants compensated for this loss with increased per capita biomass, resulting in no net change in total community biomass production along the exotic gradient (mesocosm biomass~proportion exotic, $\chi^2 = 0.01$, $p = 0.92$; fig. S2). This result, combined with the lack of direct links from NPP to ecosystem functions in the SEM (Fig. 3), suggests that the effects of exotic plant biomass on ecosystem functions were likely caused by differences in traits, leading to changes in the mean quality of plant inputs. In other words, the total quantity of biomass was the same whether communities were 100% native or 100% exotic. What differed was the relative proportion of high-quality material, which was higher in exotic-dominated communities.

Exotic plants and plants with higher SLA also supported a greater herbivore biomass (herbivore biomass~plant provenance, $\chi^2 = 26.07$, $p < 0.001$; Fig. 3 and fig. 5), a result similar to that of a previous study conducted with naturally occurring herbivores (23). Faster-growing plants with high SLA typically have tissues with proportionately lower structural

carbohydrates and, as such, are more readily mineralizable by herbivores and soil biota (13, 24). Indeed, exotic-dominated communities accumulated more herbivores (herbivore biomass~proportion exotic, $\chi^2 = 16.49$, $p < 0.001$; Fig. 2F), and this was partly driven by the higher mean SLA of these communities (Fig. 3).

Herbivore accumulation on exotic plants explains heightened decomposition rates

Previous reports have shown increases in decomposition rates as high as 117% after exotic invasion, whereas others have shown decreases [for review, see (2)]. Here, decomposition rates increased only in exotic communities in which herbivores were also present (Fig. 2C). This relationship was likely driven by the higher nutritional quality of exotic compared with native plants, with consumed biomass returned to soil by herbivores more rapidly and in a more readily mineralizable form (24). Exotic plants were more palatable, as indicated by their higher SLA, leaf N, and growth rates (24). Further, N-fixing woody and non-N-fixing herbaceous exotics had higher rates of growth despite high herbivore feeding (Fig. 3), indicating that exotics may compensate for damaged tissues more readily. Although it

is not unusual for plant–herbivore interactions to increase decomposition when herbivores feed on more palatable plants (24, 25), the accumulation of herbivores on exotic plants (particularly on non-N-fixing herbaceous exotics; Fig. 3), coupled with the high regrowth capacity of those plants, likely explains increased decomposition rates in mesocosms with a high proportion of exotic plants and herbivores added. Accumulation of successful biocontrol arthropods aimed at reducing exotic plant biomass may have similar effects in communities where they have been introduced. Our findings clarify previous idiosyncrasies from studies focused on exotic plant effects on decomposition (2), showing that exotic plants increase decomposition rates, but only when herbivores are present and able to accumulate.

Changes to C-cycling by exotic plants develops through multiple pathways

Basal and in situ respiration rates were elevated when exotic plants interacted with novel soil biota and insect herbivores. This result emerges from multiple potential pathways, the first by alteration of two bacterial taxonomic groups indicative of different C-cycling strategies. Exotic-dominated communities were associated with elevated bacterial biomass (bacterial biomass~proportion exotic, $\chi^2 = 7.54$, $p < 0.006$; Fig. 2G), particularly where N-fixing woody species were present (Fig. 3) and in communities with high mean SLA (bacterial biomass~CWSLA, $\chi^2 = 4.62$, $p < 0.03$). Herbivores depressed bacterial biomass overall (bacterial biomass~herbivore, $\chi^2 = 13.06$, $p < 0.001$; fig. S6), but also reduced the relative proportion of Gram-positive bacteria (fig. S7A), many of which are oligotrophic taxa using slow enzymatic pathways to break down recalcitrant C sources, compared with Gram-negative bacteria (often copiotrophic), which are associated with more rapid utilization of plant-based labile C sources (26–28) (Fig. 3). Aboveground herbivory can also change exudation profiles and increase plant C allocation belowground into roots (29). Thus, by increasing the amount of labile C available to soil microbes through exudates (28), herbivores may have shifted the functional attributes of the bacterial community toward those associated with the more rapid breakdown of short-term pools.

Decomposition is faster where non-N-fixing exotic plants reduce AMF

A second pathway by which exotic plants can increase soil C loss is through their reduced interactions with AMF. AMF biomass in soil decreased with exotic dominance (AMF~proportion exotic: $\chi^2 = 46.27$, $p < 0.001$; Fig. 2H) and a greater proportion of non-N-fixing plants (Fig. 3). Communities with lower AMF biomass

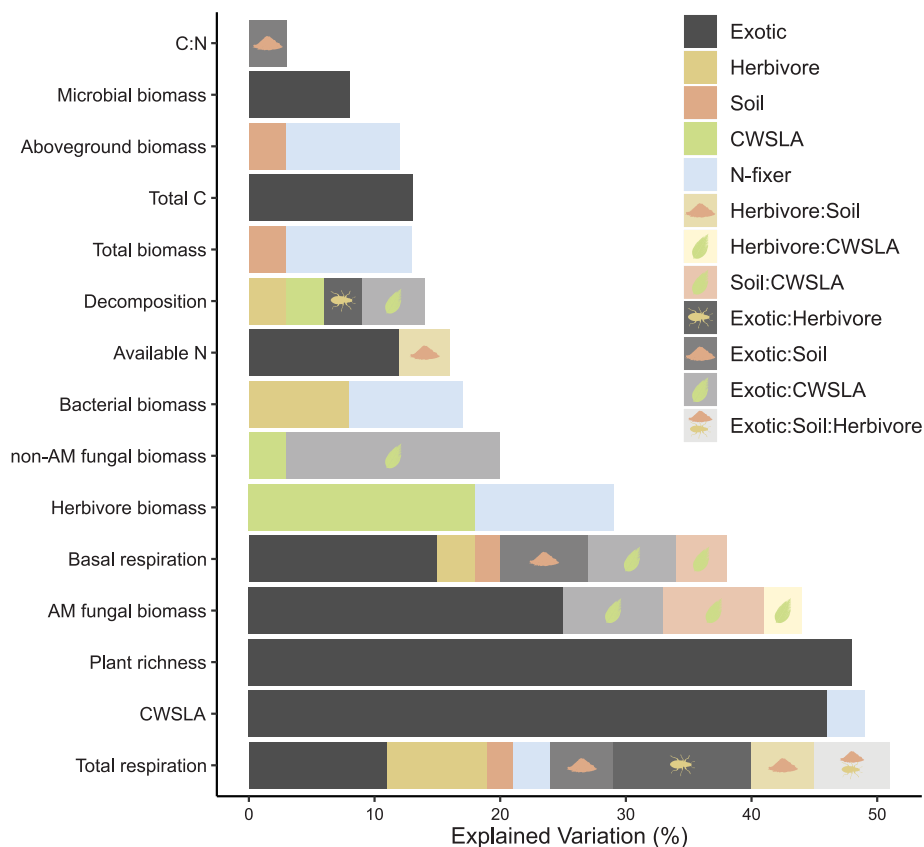


Fig. 4. Variance partitioned among treatments from hypothesized models. Bars show the percentage of variation explained by each treatment and their interactions (indicated with a colon in the key) on our measured ecosystem properties.

were associated with faster rates of decomposition (Fig. 3). AMF can stabilize organic matter in soil (30) by rapidly assimilating plant C into extensive intra- and extraradical hyphal structures (31), but much of this hyphal biomass may reside in soil as recalcitrant material and increase aggregate stabilization in soil (32). Non-AMF plants exude more labile C into soil (33, 34), which can increase mineralization of soil organic matter by cooccurring microbes (30, 33, 34). Indeed, suppression of AMF has accelerated soil respiration in previous studies (35). By contrast, ectomycorrhizal plants stimulate C mineralization in soils (33). However, we did not observe increases in C turnover in communities with ectomycorrhizal plants, likely because communities containing these species typically contained only a single ectomycorrhizal individual and made up a relatively small proportion of biomass (~6% where they occurred).

Although AMF are ubiquitous and associate with most land plants, there is wide variation in the degree to which plants invest in AMF (36), which may be species specific (36) and differ depending on provenance (37) or environmental conditions (38). A meta-analysis has shown that native plants have lower AMF colonization when grown near or after exotic plants, but provenance did not predict differences in AMF abundance among exotic plants (37). Here, exotic non-N-fixing plants caused the biggest declines in plant richness (Fig. 3), which may have reduced the availability of potentially better hosts for AMF. N-fixing plants did not reduce AMF biomass, likely reflecting their high reliance on AMF to meet the high phosphorus demand of N fixation (39). AMF were also strongly reduced by non-N-fixing woody plants (Fig. 3), possibly reflective of suppression by ectomycorrhizal plants and associated fungi (40). Finally, direct suppression of AMF hyphae has been observed in soils with a high abundance of Gram-negative bacteria (41), which were elevated in our exotic-dominated communities. Thus, increases in potentially antagonistic bacteria or other biota may explain the concomitant reductions in AMF biomass in exotic-dominated communities, or reduced biomass may reflect lower investment in AMF by the exotic plants grown here. AMF can buffer C emissions as long as decomposition rates remain low (42, 43), suggesting that decreases in AMF biomass under exotic plants represents another avenue by which C dynamics shift after invasion.

Increased microbial respiration rates in exotic-dominated communities grown in soils with depressed levels of species-specific soil biota (i.e., away soils) (Fig. 2B) were also likely driven by functional changes to bacterial communities. Plants grown in soils including species-specific soil biota (i.e., home soils) had a higher proportion of Gram-positive

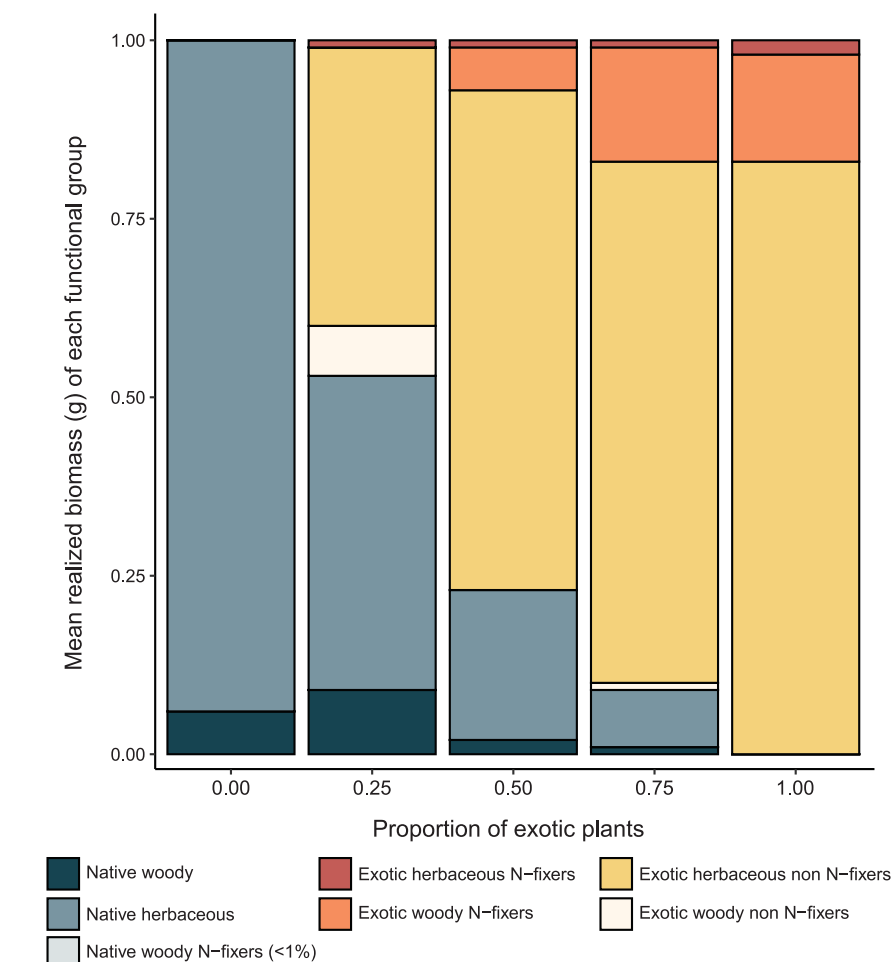


Fig. 5. Mean proportion of plant biomass realized relative to proportion planted by provenance and functional group. Bars represent the mean realized proportion of biomass for exotic and native plants of each plant functional group and provenance along the planted exotic gradient.

bacteria (Gram-positive bacteria~soil, $\chi^2 = 3.94$, $p < 0.047$; fig. S7B), a predominantly oligotrophic group containing many important, often specialized plant pathogens and recalcitrant litter specialists (27). Although we cannot determine whether soil biota were beneficial or harmful to plants, community-level reductions in plant biomass in home soil support the notion of more highly specialized pathogen loads or poor-quality mutualists (mesocosm plant biomass~soil, $\chi^2 = 8.80$, $p < 0.003$; Fig. 4 and fig. S4). Regardless of their effects on live plants, oligotrophic taxa often have slower growth and respiratory loss rates (27, 28) and may have slowed respiratory losses in home soils. In fact, increases in Gram-positive bacteria were associated with higher soil organic matter (Fig. 3), likely explaining the higher C:N in home soils. By contrast, the higher basal respiration rates in away soils likely reflect the relatively rapid breakdown of labile plant material by more prevalent copiotrophic taxa. This provides a temporal aspect to invasion effects on C dy-

namics. If home soils (containing more specialist soil biota) represent more established invasions and away soils newer invasions (containing no specialist soil biota), then we can expect faster turnover of recently assimilated C in new invasions, shifting to slower turnover and greater storage of longer-term C pools over time as specialist bacteria increase in abundance.

Our results could inform projects aimed at using nature-based solutions to fight climate change, such as the Trillion Trees Program (www.trilliontrees.org/). Fast-growing exotic plants may be ideally suited for plantings when the goal is to sequester C aboveground, but exotic plants may have more variable effects belowground compared with native plants. Currently, N-fixing trees from the genus *Acacia* are among the most commonly planted plantation trees globally (44). Results from our experiment, which included *A. dealbata*, suggest that these types of trees may indeed increase soil C stocks, particularly in soils previously occupied by those species (i.e., home soils).

through their positive effects on oligotrophic, specialist soil biota. However, these gains may be reduced over time through respiration associated with increases in copiotrophic taxa such as those that we have observed (Fig. 3), particularly in new or frequently harvested plantations. Our results also suggest that soil C stability may vary depending on whether exotic plants accumulate herbivores and/or fungal saprotrophs regardless of woodiness.

The pathways by which N-fixing versus non-N-fixing exotics affect ecosystem properties related to C-cycling appear to differ, and their relative effects differed between woody and herbaceous plants. N-fixing woody exotics disproportionately affected ecosystem functions relative to their proportional biomass in communities (Fig. 5), which has also been observed in other systems (45). Despite their low richness, the five exotic N-fixing woody species with high survival drove the biggest increases in fungal and bacterial biomass (Fig. 3), likely explaining much of the increased microbial respiration rates along the exotic gradient (Fig. 3). Non-N-fixing species contributed to significant decreases in AMF biomass, which may be driven by their exploitative life history strategies [i.e., fast growth (46) and high SLA; Figs. 2 and 3]. Herbaceous N-fixing exotics had no detectable effects in their communities, whereas non-N-fixing woody species had few (but strong) effects, perhaps because these species did not reach maturity, whereas many of the herbaceous plants grew quite large and usually flowered. Thus, exotic effects on ecosystem functions consistently result from differences in the way that they interact with other organisms in their environment compared with their native counterparts. Our experiment reveals multiple pathways through which these effects occur, influenced in different ways by various plant traits, including leaf traits, plant functional group (i.e., woody or herbaceous, N-fixing or not), and their typical microbial associations.

Taken together, these results suggest that shifts in fungal and bacterial function in exotic-dominated communities with herbivores and more generalist soil biota may underpin increased respiration rates through more rapid microbial mineralization of organic matter. Soil organic matter is primarily derived from plants and developed and turned over through interactions among soil microbiota (47), the abiotic environment (48), and living plant root exudates (49), which can differ substantially depending on plant composition (50). Indeed, our analyses indicate that it is not simply the large increases in exotic plant biomass that drive changes to C-cycling, but also differences in the types of inputs compared with natives. Exotic N-fixers drive increases in faster-cycling bacterial taxa, whereas high-quality exotic leaf tissues promote biomass of herbivores that

reduce slower-cycling relative to faster-cycling taxa (Fig. 3 and fig. S7).

Understanding the factors regulating ecosystem responses in invaded ecosystems is complicated by the myriad of factors that influence C gains and losses. Our manipulative experiment showed strong interdependence among three leading hypotheses used to explain drivers of exotic plant impacts on ecosystem properties related to C-cycling. Exotic plants, which make up the greatest proportion of biomass in many invaded communities, have traits (e.g., high SLA and N_{mass}) that influence the way that they interact with herbivores and soil biota to drive ecosystem properties. Our results show that exotic and native plants differ in common bacterial and fungal pathways of C storage and release, and these differences are driven by trait differences such as SLA, woodiness, and the ability to fix N. Exotic plants, particularly woody species that fix N, increase bacterial and non-AMF biomass associated with storage and breakdown of C pools. By contrast, non-N-fixing exotic herbs substantially reduce AMF biomass, which accelerates soil C loss compared with communities dominated by native plants. In addition to their effects on fungi and bacteria, herbaceous exotic plants increase herbivore biomass, which we linked to increases in decomposition, an indicator of C-cycling. Our study provides an unprecedented community-scale understanding of how the transformation of ecosystems by exotic species depends on complex species interactions, emphasizing the need for whole-ecosystem approaches.

REFERENCES AND NOTES

- J. G. Ehrenfeld, *Ecosystems* (N. Y.) **6**, 503–523 (2003).
- C. Liao et al., *New Phytol.* **177**, 706–714 (2008).
- D. A. Peltzer, R. B. Allen, G. M. Lovett, D. Whitehead, D. A. Wardle, *Glob. Change Biol.* **16**, 732–746 (2010).
- J. G. Ehrenfeld, *Annu. Rev. Ecol. Syst.* **41**, 59–80 (2010).
- W. H. van der Putten, J. N. Klironomos, D. A. Wardle, *ISME J.* **1**, 28–37 (2007).
- B. E. Wolfe, J. N. Klironomos, *Bioscience* **55**, 477–487 (2005).
- M. E. Rout, R. M. Callaway, *Science* **324**, 734–735 (2009).
- M. van Kleunen, E. Weber, M. Fischer, *Ecol. Lett.* **13**, 235–245 (2010).
- A. A. Agrawal et al., *Ecology* **86**, 2979–2989 (2005).
- See the supplementary materials and methods.
- J. D. Bever, K. M. Westover, J. Antonovics, *J. Ecol.* **85**, 561–573 (1997).
- C. A. Johnston et al., *Front. Ecol. Environ.* **2**, 522–528 (2004).
- I. J. Wright et al., *Nature* **428**, 821–827 (2004).
- H. Poorter, J. R. Evans, *Oecologia* **116**, 26–37 (1998).
- E. Laliberté, J. M. Tylianakis, *Ecology* **93**, 145–155 (2012).
- R. D. Bardgett, L. Mommer, F. T. De Vries, *Trends Ecol. Evol.* **29**, 692–699 (2014).
- K. N. Suding et al., *Glob. Change Biol.* **14**, 1125–1140 (2008).
- M. Tamura, N. Tharayil, *New Phytol.* **203**, 110–124 (2014).
- J. L. Funk, P. M. Vitousek, *Nature* **446**, 1079–1081 (2007).
- J. M. Levine et al., *Proc. Biol. Sci.* **270**, 775–781 (2003).
- C. C. Daehler, *Annu. Rev. Ecol. Syst.* **34**, 183–211 (2003).
- M. Vilà et al., *Ecol. Lett.* **14**, 702–708 (2011).
- A. A. Agrawal, P. M. Kotanen, *Ecol. Lett.* **6**, 712–715 (2003).

- D. J. Augustine, S. J. McNaughton, *J. Wildl. Manage.* **62**, 1165–1183 (1998).
- D. A. Wardle et al., *Science* **304**, 1629–1633 (2004).
- C. Kramer, G. Gleixner, *Soil Biol. Biochem.* **40**, 425–433 (2008).
- N. Fierer, M. A. Bradford, R. B. Jackson, *Ecology* **88**, 1354–1364 (2007).
- N. Farin et al., *Soil Biol. Biochem.* **128**, 111–114 (2019).
- J. N. Holland, W. Cheng, D. A. Crossley Jr., *Oecologia* **107**, 87–94 (1996).
- D. L. Jones, A. Hodge, Y. Kuzyakov, *New Phytol.* **163**, 459–480 (2004).
- P. L. Staddon, C. B. Ramsey, N. Ostle, P. Ineson, A. H. Fitter, *Science* **300**, 1138–1140 (2003).
- Y. G. Zhu, R. M. Miller, *Trends Plant Sci.* **8**, 407–409 (2003).
- R. P. Phillips, T. J. Fahey, *Ecology* **87**, 1302–1313 (2006).
- E. R. Zorostek, D. Dragoni, Z. A. Brown, R. P. Phillips, *New Phytol.* **206**, 1274–1282 (2015).
- B. Zhang et al., *Sci. Rep.* **6**, 19990 (2016).
- G. W. Wilson, D. C. Hartnett, *Am. J. Bot.* **85**, 1732–1738 (1998).
- R. A. Bunn, P. W. Ramsey, Y. Lekberg, *J. Ecol.* **103**, 1547–1556 (2015).
- N. C. Johnson, J. H. Graham, F. A. Smith, *New Phytol.* **135**, 575–585 (1997).
- P. Bonfante, I.-A. Anca, *Annu. Rev. Microbiol.* **63**, 363–383 (2009).
- K. M. Becklin, M. L. Pallo, C. Galen, *J. Ecol.* **100**, 343–351 (2012).
- N. B. Svenningsson et al., *ISME J.* **12**, 1296–1307 (2018).
- E. F. Leifheit, E. Verbruggen, M. C. Rillig, *Soil Biol. Biochem.* **81**, 323–328 (2015).
- E. Verbruggen, J. Jansa, E. C. Hammer, M. C. Rillig, *J. Ecol.* **104**, 261–269 (2015).
- M. Dodet, C. Collet, *Biol. Invasions* **14**, 1765–1778 (2012).
- J. M. Craine, W. G. Lee, S. Walker, “The context of plant invasions in New Zealand: Evolutionary history and novel niches” in *Biological Invasions in New Zealand* (Springer, 2006), pp. 167–177.
- L. Koziol, J. D. Bever, *Ecology* **96**, 1768–1774 (2015).
- D. W. Hopkins, E. G. Gregorich, *Biol. Div. Func. Soil* **2005**, 57–80 (2005).
- J. F. Dormaar, *Biol. Fertil. Soils* **10**, 121–126 (1990).
- J. A. Bird, D. J. Herman, M. K. Firestone, *Soil Biol. Biochem.* **43**, 718–725 (2011).
- F. Z. Haichar et al., *ISME J.* **2**, 1221–1230 (2008).
- L. P. Waller et al., Data for: Biotic interactions drive ecosystem responses to exotic plant invaders, Dryad (2020); <https://doi.org/10.5061/dryad.wwp2gm5gb>.

ACKNOWLEDGMENTS

We thank D. Wardle and E. T. Kiers for valuable comments and J. Allen, N. Allen, A. Barrington, G. Boitt, J. Breitmeyer, J. Bufford, J. Burson, D. Conder, R. Cresswell, D. Dash, L. Dickie, B. Fairbrass, C. Ferguson, B. Ganley, T. Glare, A. Holyoake, D. Jack, C. Johns, J. Johns, E. Jones, B. Kwan, S. Larsen, H. Lea, I. Luxford, F. Martoni, L. Meachen, A. McKinnon, G. McSweeney, M. O’Callaghan, A. Puértolas, B. Richards, S. Richardson, H. Ridgeway, R. Scott, M.-R. Shadboit, K. Slattery, R. Vardy, R. Wainer, D. Waller, M. Waller, A. Wakelin, and S. Wilson for technical support. **Funding:** This research was primarily supported by Tertiary Education Commission funding to the BioProtection Research Centre. **Author contributions:** Conceptualization: W.J.A., L.M.C., I.A.D., J.M.T., L.P.W.; Funding acquisition: B.I.P.B., L.M.C., I.A.D., J.M.T., S.A.W.; Investigation: W.J.A., F.M.F., J.E.H., N.K., K.H.O., G.S.S., L.P.W.; Data analysis: W.J.A., I.A.D., J.M.T., L.P.W.; Supervision: B.I.P.B., L.M.C., I.A.D., J.M.T.; Writing original draft: L.P.W.; Writing, review, and editing: all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The data reported here have been deposited in Dryad (51).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6494/967/suppl/DC1
Materials and Methods

Figs. S1 to S7
Tables S1 to S7
References (52–67)

[View/request a protocol for this paper from Bio-protocol.](#)

2 December 2019; accepted 20 April 2020
10.1126/science.aba2225

CANCER

The human tumor microbiome is composed of tumor type-specific intracellular bacteria

Deborah Nejman^{1*}, Ilana Livvyatan^{1,2*}, Garold Fuks^{3*}, Nancy Gavert¹, Yaara Zwang¹, Leore T. Geller¹, Aviva Rotter-Maskowitz¹, Roi Weiser^{4,5}, Giuseppe Malle¹, Elinor Gigi¹, Arnon Meltser¹, Gavin M. Douglas⁶, Iris Kamer⁷, Vancheswaran Gopalakrishnan^{8†}, Tali Dadosh⁹, Smadar Levin-Zaidman⁹, Sofia Avnet¹⁰, Tehila Atlan¹¹, Zachary A. Cooper¹², Reetakshi Arora⁸, Alexandria P. Cogdill¹³, Md Abdul Wadud Khan⁸, Gabriel Ologun⁸, Yuval Bussi^{1,2,14}, Adina Weinberger^{1,2}, Maya Lotan-Pompan^{1,2}, Ofra Golani¹⁵, Gili Perry¹⁶, Merav Rokah¹⁷, Keren Bahar-Shany¹⁶, Elisa A. Rozeman¹⁸, Christian U. Blank¹⁸, Anat Ronai¹⁹, Ron Shaoul¹⁹, Amnon Amit^{20,21}, Tatiana Dorfman^{22,23}, Ran Kremer²⁴, Zvi R. Cohen^{5,25}, Sagi Harnov^{5,26}, Tali Siegal²⁷, Einav Yehuda-Shnaidman²⁸, Einav Nili Gal-Yam²⁹, Hagit Shapira²⁸, Nicola Baldini^{10,30}, Morgan G. I. Langille^{6,31}, Alon Ben-Nun^{5,17}, Bella Kaufman^{5,7}, Aviram Nissan³², Talia Golan^{5,7}, Maya Dadiani¹⁶, Keren Levanon^{5,16}, Jair Bar^{5,7}, Shlomit Yust-Katz^{5,27}, Iris Barshack^{5,33}, Daniel S. Peeper³⁴, Dan J. Raz³⁵, Eran Segal^{1,2}, Jennifer A. Wargo^{8,13}, Judith Sandbank²⁸, Noam Shental^{36†}, Ravid Straussman^{1†§}

Bacteria were first detected in human tumors more than 100 years ago, but the characterization of the tumor microbiome has remained challenging because of its low biomass. We undertook a comprehensive analysis of the tumor microbiome, studying 1526 tumors and their adjacent normal tissues across seven cancer types, including breast, lung, ovary, pancreas, melanoma, bone, and brain tumors. We found that each tumor type has a distinct microbiome composition and that breast cancer has a particularly rich and diverse microbiome. The intratumor bacteria are mostly intracellular and are present in both cancer and immune cells. We also noted correlations between intratumor bacteria or their predicted functions with tumor types and subtypes, patients' smoking status, and the response to immunotherapy.

More than 16% of cancer incidence worldwide has been attributed to infectious agents (1). Intratumor bacteria have been reported in many tumor types (2–19), but these bacteria have not been characterized in a comprehensive manner (20). The gut microbiome has been shown to have multiple effects on tumor biology, such as the transformation process, tumor progression, and the response to anticancer therapies including immunotherapy (21–29). Thus, characterization of the tumor microbiome may be an essential step in unraveling the effects that tumor bacteria have on different cancer hallmarks.

Bacterial DNA, RNA, and lipopolysaccharide are present in many human solid tumors

Because the tumor microbiome has a relatively low biomass, contamination of the tumor

samples with bacteria or bacterial DNA can be problematic (30, 31). Therefore, it is critical to include multiple measures to avoid, or at least detect, any possible contamination in the process of profiling the tumor microbiome (supplementary note) (32, 33). For next-generation sequencing applications, it is also important to use mechanical tissue shearing in the DNA isolation protocol to ensure the complete recovery of DNA from within the cell wall of Gram-positive bacteria—a step not included in most standard DNA isolation protocols (6, 7). To characterize and visualize intratumor bacteria, we applied an extensive combination of methods to a large cohort of solid human tumor samples to detect bacterial DNA, RNA, and bacterial outer membrane or cell wall components.

We focused on seven solid tumor types that represent either common cancer types or cancer

types for which the tumor microbiome is largely unknown, such as melanoma, bone, and brain tumors (Fig. 1A). To address laboratory-borne contaminants, we introduced 643 negative controls that were processed with the samples, including 437 DNA extraction controls and 206 polymerase chain reaction (PCR) no-template controls (NTCs). To address contamination that might have occurred before the samples reached our laboratory, we also included 168 paraffin-only samples taken from the margins of the paraffin blocks (without tissue) that were used in the study (Fig. 1A).

Overall, we profiled 1010 tumor samples and 516 normal samples, including normal adjacent tissues (NATs) from the same patients (Fig. 1A and table S1). In the case of ovarian cancer, our normal samples came from the ovaries or uteruses of the patients or from normal fallopian tube fimbria of unmatched healthy subjects (tables S1 and S2). To quantify bacterial DNA, we used a real-time quantitative PCR (qPCR) assay with universal primers 967F and 1064R specific for the bacterial ribosomal 16S gene [16S rDNA (ribosomal DNA)] (34). Levels of bacterial DNA in all tumor types were significantly higher than those found in both DNA extraction and paraffin controls (Fig. 1B; P value $<10^{-10}$ for each tumor type, Wilcoxon rank sum test). We found that different cancer types vary in the proportion of tumors that are positive for bacterial DNA, ranging from only 14.3% in melanoma to >60% in breast, pancreatic, and bone tumors. Bacterial DNA was also detected in solid tumors that have no direct connection with the external environment, such as ovarian cancer, glioblastoma multiforme (GBM), and bone cancer.

To validate the presence of bacteria in human tumors, we stained >400 additional tumors (not related to the 1526 samples described above), representing six of our seven profiled tumor types, for the presence of bacteria. We conducted immunohistochemistry (IHC) using antibodies against bacterial lipopolysaccharide (LPS) and lipoteichoic acid (LTA) to detect Gram-negative and Gram-positive bacteria, respectively (35, 36). We also used RNA fluorescence in situ hybridization (FISH), with a

¹Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel. ²Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel.

³Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel. ⁴Division of Surgery, Tel-Aviv Sourasky Medical Center, Tel-Aviv, Israel. ⁵Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel. ⁶Department of Microbiology and Immunology, Dalhousie University, Halifax, NS, Canada. ⁷Institute of Oncology, Sheba Medical Center, Ramat Gan, Israel.

⁸Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁹Department of Chemical Research Support, Weizmann Institute of Science, Rehovot, Israel. ¹⁰Orthopaedic Pathophysiology and Regenerative Unit, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy. ¹¹Department of Bioinformatics, Jerusalem College of Technology, Jerusalem, Israel.

¹²Translational Medicine, Oncology R&D, AstraZeneca, Gaithersburg, MD, USA. ¹³Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ¹⁴Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot, Israel. ¹⁵Department of Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot, Israel. ¹⁶Cancer Research Center, Sheba Medical Center, Ramat Gan, Israel. ¹⁷Department of Thoracic Surgery, Sheba Medical Center, Ramat Gan, Israel. ¹⁸Department of Medical Oncology and Division of Molecular Oncology and Immunology, Netherlands Cancer Institute, Amsterdam, Netherlands. ¹⁹Pediatric Gastroenterology Institute, Rambam Medical Center, Haifa, Israel. ²⁰Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel. ²¹Department of Obstetrics and Gynecology, Rambam Health Care Campus, Haifa, Israel. ²²Division of General Surgery, Rambam Health Care Campus, Haifa, Israel. ²³Ambulatory and Breast Surgery Service, Rambam Health Care Campus, Haifa, Israel. ²⁴Department of Thoracic Surgery, Rambam Health Care Campus, Haifa, Israel. ²⁵Department of Neurosurgery, Sheba Medical Center, Ramat Gan, Israel. ²⁶Department of Neurosurgery, Rabin Medical Center, Beilinson Hospital, Petach Tikva, Israel. ²⁷Neuro-Oncology Unit, Rabin Medical Center, Beilinson Hospital, Petach Tikva, Israel. ²⁸Institute of Pathology, Megalab, Maccabi Healthcare Services, Rehovot, Israel. ²⁹Breast Oncology Institute, Sheba Medical Center, Ramat Gan, Israel. ³⁰Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy. ³¹Department of Pharmacology, Dalhousie University, Halifax, NS, Canada. ³²Department of Surgical Oncology (Surgery C), Sheba Medical Center, Ramat Gan, Israel. ³³Department of Pathology, Sheba Medical Center, Ramat Gan, Israel. ³⁴Division of Molecular Oncology & Immunology, Netherlands Cancer Institute, Amsterdam, Netherlands. ³⁵Division of Thoracic Surgery, City of Hope Medical Center, Duarte, CA, USA. ³⁶Department of Mathematics and Computer Science, The Open University of Israel, Ra'anana, Israel.

*These authors contributed equally to this work. †Present address: AstraZeneca, Gaithersburg, MD, USA. ‡These authors contributed equally to this work.

§Corresponding author. Email: ravidst@weizmann.ac.il

universal probe against bacterial 16S ribosomal RNA (rRNA), to detect bacterial RNA in these tumors (37). To control for nonspecific staining, IHC-negative controls (no primary antibody) and FISH-negative controls (nonspecific complement probe) were also applied to the samples (figs. S1 and S2). Bacterial LPS and 16S rRNA were frequently detected in all tumor types (Fig. 1C) and demonstrated a similar spatial distribution (Fig. 1D and fig. S3). LTA was detected mainly in melanomas and was largely absent in other tumor types.

Generally, more tumors were found to be positive for bacteria using visualization methods than by using qPCR. This disparity may be because of some limitation in the sensitivity of our qPCR assay, or it might be caused by our strict cutoff for confirming a sample as positive.

Intratumor bacteria are mostly intracellular and are present in both cancer and immune cells

Pathological examination of tumor cores indicated that LPS and bacterial 16S rRNA were

localized mainly in cancer cells and immune cells (Fig. 2A and fig. S4). In cancer cells, bacterial 16S rRNA was detected mostly in the cytoplasm, whereas LPS staining was associated with both the cytoplasm and the nucleus (Fig. 2B and fig. S5). CD45-positive leukocytes generally exhibited a stronger cytoplasmic bacterial staining by 16S rRNA staining than that exhibited by cancer cells (Fig. 2C and fig. S6A). LTA-positive bacteria were almost exclusively found in macrophages, as detected by hematoxylin and eosin (H&E) staining and verified by immunofluorescence (IF) for CD68 (Fig. 2D and fig. S6B). LTA was rarely detected in cancer cells or in CD45+/CD68– immune cells (Fig. 2). Although the intensity of bacterial LPS and LTA staining was very strong in CD45+/CD68+ cells, bacterial 16S rRNA was only rarely found in macrophages by FISH (Fig. 2, A and D, and figs. S4 and S6). This discrepancy may reflect macrophage ingestion of bacterial components rather than live bacteria, or it may result from the accumulation of LPS and LTA in macrophages long after the bacteria have been phagocytized and processed by the macrophages. It has been previously demonstrated that the processing of bacterial LPS by macrophages is very slow; therefore, LPS can be found in these cells months after the bacteria were ingested and processed (38).

To further verify the presence of bacteria inside cancer cells, we performed correlative light and electron microscopy (CLEM) (39, 40) on four human breast tumors that were positive for bacterial LPS and 16S rRNA (fig. S7). Combined LPS fluorescence staining and transmission electron microscopy (TEM) imaging of the same cells clearly demonstrated the intracellular localization of bacteria in all four tumors (Fig. 2E and fig. S7). In many cases, the bacteria were found in close proximity to the nuclear membrane. Because we did not detect intranuclear bacteria by TEM, we suspect that the appearance of LPS nuclear localization in some tumors represents the staining of cytoplasmic perinuclear bacteria.

Whereas bacterial 16S rRNA FISH staining appeared as a diffused signal inside cells, typical bacterial rods or cocci were only rarely detected (in 3 of 426 cores). This observation, combined with the fact that no cell wall polymer LTA was detected in cancer cells—despite the detection of many Gram-positive bacteria in human tumors by 16S rDNA sequencing—suggests that bacteria in tumor cells may have altered their envelope, perhaps leading to a cell wall-deficient state, akin to L-forms (41). Cell wall-deficient bacteria are known to be found exclusively inside cells, where their morphology transforms into less-defined structures of highly variable sizes and shapes (42, 43). Our TEM images also suggest that many of the

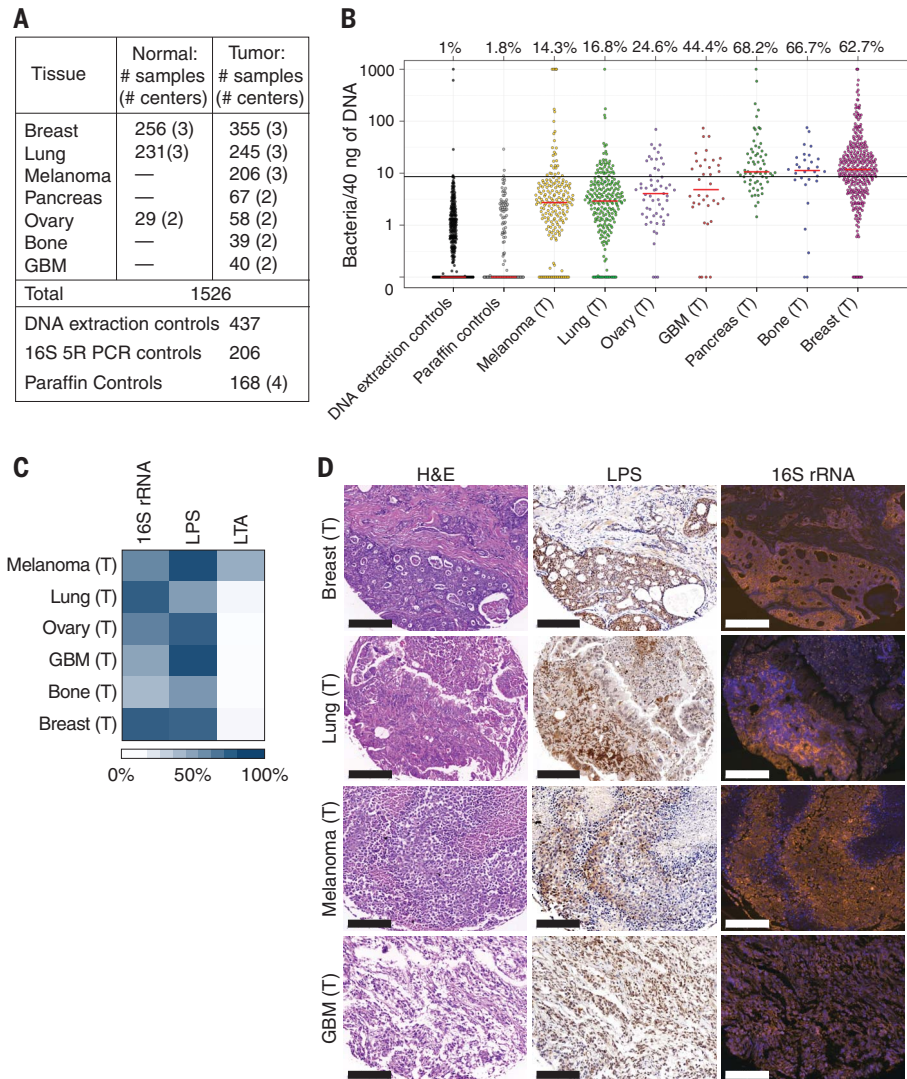


Fig. 1. Bacterial components are detected in human tumors. (A) Number of human samples analyzed in the study. Normal samples include both normal tissue samples and normal adjacent tissue (NAT) to tumor samples, as detailed in table S1. Dashes indicate data not available. GBM, glioblastoma multiforme. (B) The presence of bacterial DNA in human tumors was assessed by bacterial 16S rDNA qPCR. A calibration curve, generated by spiking bacterial DNA into human DNA, was used to estimate bacterial load, which was then normalized against batch-specific qPCR NTCs. Values were floored to 0.1. Red bars represent the median. The proportion of samples of each cancer type that had more bacteria than the 99th percentile of the DNA extraction control samples (black bar) is depicted above each cancer type. (C) Heatmap representing the proportion of tumors that stained positively for 16S rRNA, LPS, or LTA. $n = 40$ to 101 tissue cores per tumor type. (D) Consecutive slices from four human tumor types were stained with H&E, anti-LPS antibody (LPS), or with FISH probes against bacterial 16S rRNA. Scale bars, 200 μ m. The letter (T) indicates samples originating from tumors.

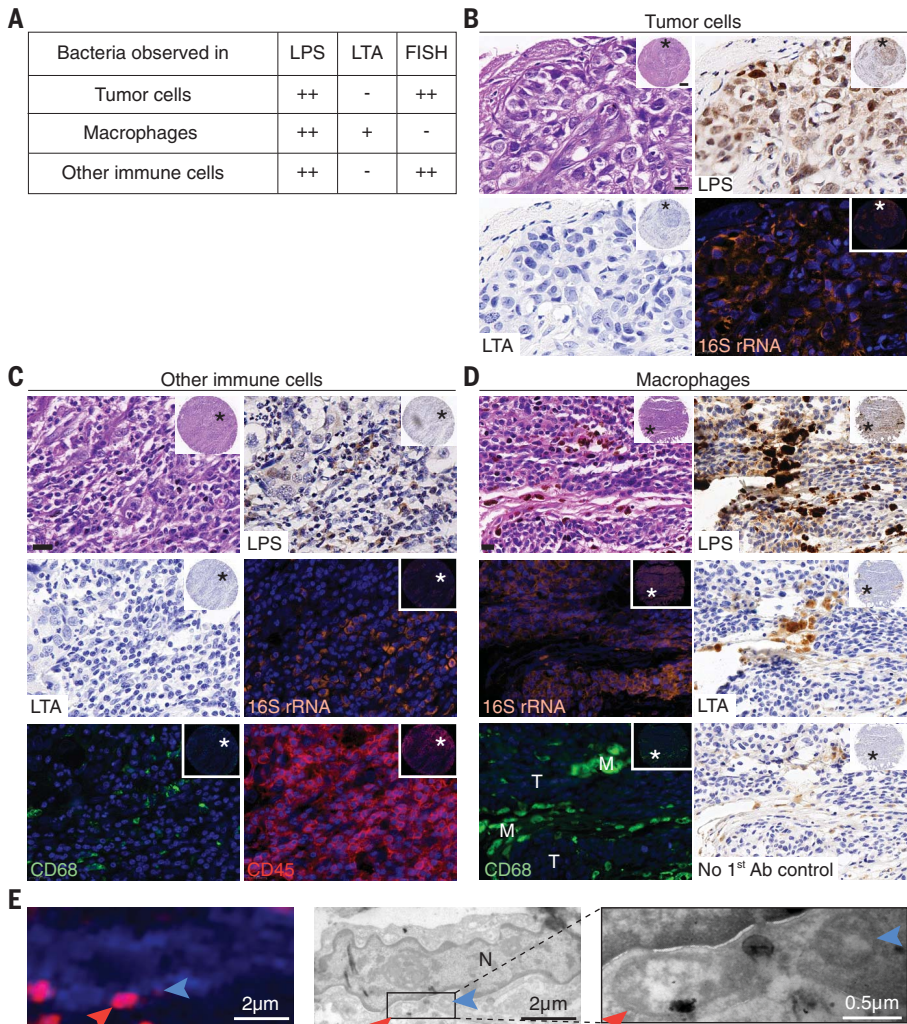


Fig. 2. Intratumor bacteria are found inside both cancer and immune cells. (A) Summary of the staining patterns of LPS, LTA, and bacterial 16S rRNA in different cell types across 459, 427, and 354 tumor cores, respectively. CD45+/CD68+ cells are referred to as macrophages; CD45+/CD68- cells are referred to as other immune cells. (B to D) Representative cores are shown demonstrating the different staining patterns in human tumors. Asterisks mark the region that was selected for higher magnification. (B) Bacterial LPS and 16S rRNA are demonstrated in breast cancer cells. (C) Bacterial LPS and 16S rRNA are demonstrated in CD45+/CD68- cells of a highly inflamed breast tumor. (D) A melanoma tumor demonstrating typical staining of macrophage-associated bacteria (M), with positive LPS and LTA staining but no 16S rRNA staining. Nearby tumor cells (T) show the typical LPS and 16S rRNA staining, with negative LTA staining. Each inset demonstrates a low magnification of the entire core. Scale bars in high-magnification images, 20 μ m. (E) CLEM demonstrates intracellular bacteria in human breast cancer. IF image shows DAPI in blue and LPS in red. Two bacteria are marked with arrows. TEM images of the same cell are shown in grayscale. High-magnification image of the boxed area is shown on the right. The letter N marks the cell nucleus.

intracellular bacteria lack a cell wall (Fig. 2E and fig. S7).

The microbiome of breast tumors is richer and more diverse than that of other tumor types

To characterize the intratumor microbiome, we developed a multiplexed 16S rDNA sequencing protocol that amplifies five short regions along the 16S rRNA gene: the 5R 16S rDNA sequencing method (Fig. 3A). By amplifying 68% of the bacterial 16S rRNA gene

using short amplicons, this method increases the coverage and resolution of the detection of bacterial species compared with the widely used V4 or V3-V4 amplification (fig. S8). Moreover, it can be applied to partially degraded DNA originating from formalin-fixed paraffin-embedded tumors. Reads from 1526 samples and 811 negative controls (DNA extraction controls, 16S 5R PCR controls, and paraffin controls) were computationally combined into long amplicons, using Short Multiple Regions

Framework (SMURF) (44) and the Greengenes database as a reference. To improve taxonomic assignment, we used the Ribosomal Database Project (RDP) classifier to augment the Greengenes database by assigning a species-level taxonomy to 380,000 bacterial 16S rRNA sequences that originally lacked such taxonomy (45) (table S3 and materials and methods). Thirty-nine samples and 10 controls that had fewer than 1000 normalized reads were discarded from further analysis (materials and methods).

Overall, we detected 9190 bacterial species across the different tumor or normal tissue types (Fig. 3B and table S2). Because some of these species may represent contamination of the samples, we applied a strict set of six filters to control for potential sources of contamination. To account for the most frequent general contaminants, filter 1 removed 167 bacterial species that were detected in >7.5% of our DNA extraction and NTC negative control samples or in the paraffin controls. This threshold demarcates the transition between most of the species that are absent or very rarely present in controls and the species that appear much more commonly in controls (fig. S9). We then applied three filters to control for batch effects that originate from DNA extraction, PCR amplification, or sequencing lane using hundreds of negative controls as a background for laboratory-borne contamination (filters 2 to 4). Filters 5 and 6 were added to control for contamination that might have been introduced to the samples before their processing in the laboratory. Filter 5 uses paraffin-only samples (without tissue) from the margins of the same paraffin blocks that were used in the study to control for contamination in the process of preparing and storing the paraffin blocks. Lastly, to account for other potential sources of medical center-specific contamination, filter 6 excluded bacteria that were not significantly enriched in a specific tumor type across multiple medical centers. Only bacteria that passed all six filters in a specific cancer type or its NAT were considered to be hits that are present in this cancer or NAT condition (Fig. 3B, table S4, materials and methods, and supplementary note).

We found that breast tumors had a richer and more diverse microbiome than all other tumor types tested (P value $<10^{-15}$ for each tumor type, Wilcoxon rank sum test; Fig. 3, C and D, and figs. S10 and S11). An average of 16.4 bacterial species were detected in any single breast tumor sample, whereas the average was <9 in all other tumor types (P value $<10^{-17}$ for each tumor type, Wilcoxon rank sum test; Fig. 3E and fig. S11). We also found that bacterial load and richness were higher in the breast tumor samples than those found in normal breast samples from healthy subjects.

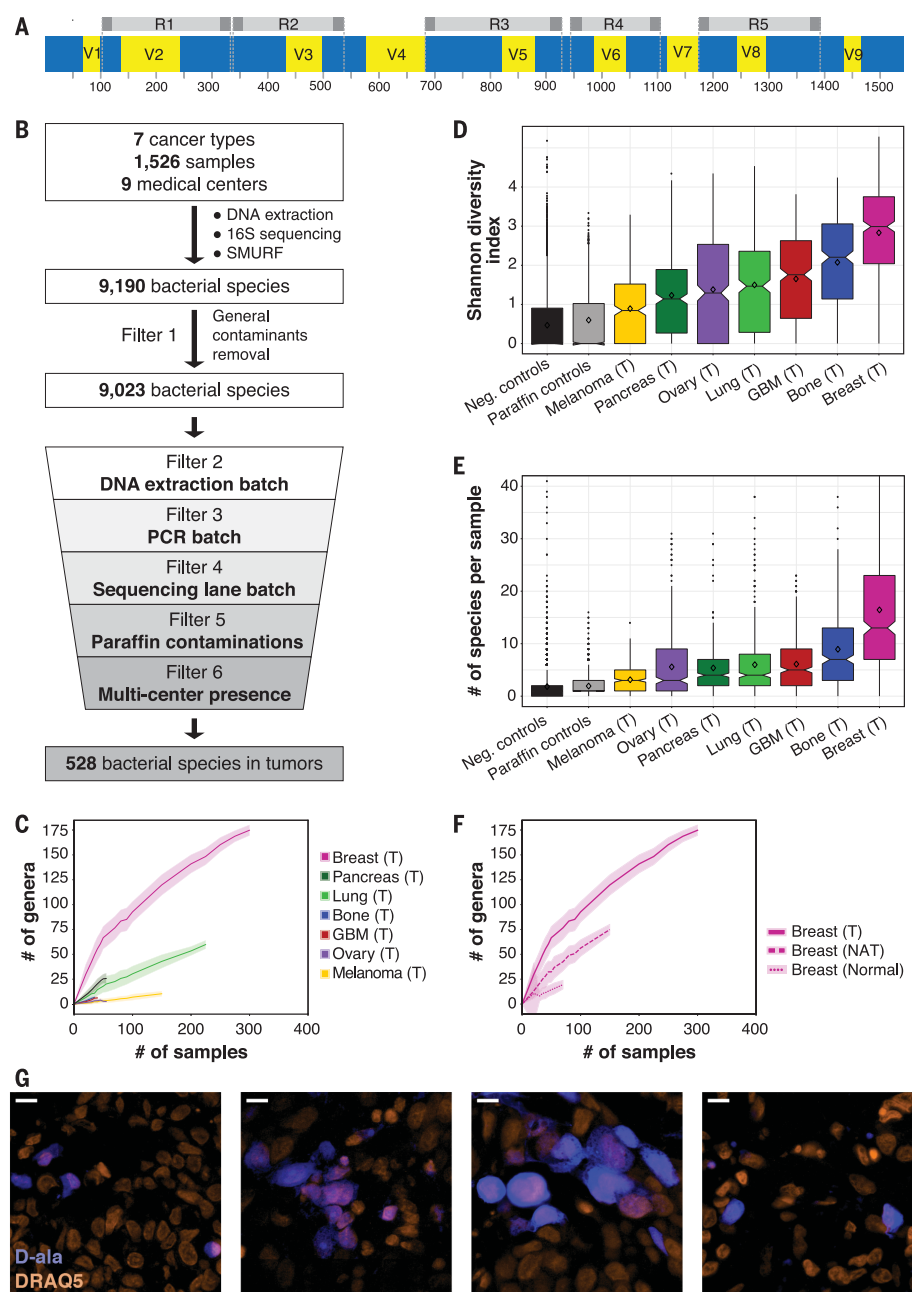


Fig. 3. The microbiome of breast tumors is richer and more diverse than that of other tumor types.

(A) Graphic representation of the bacterial 16S rRNA gene with its conserved (blue) and variable (yellow) regions. The sequence from *Escherichia coli* K-12 strain MG1655 was used as a reference sequence. The five amplicons of the multiplexed 5R PCR method are depicted in gray. (B) Schematic representation of the analysis pipeline applied to 16S rDNA sequencing data. (C) Rarefaction plots showing the number of bacterial genera that passed all filters in the different tumor types per number of tumor samples that were selected for the analysis. Light color shading represents confidence intervals based on 100 random subsamplings for each number of tumor samples that was analyzed. (D) Box blot of Shannon diversity indexes of all samples, segregated by tumor type. Neg., negative. (E) Box blot of the numbers of bacterial species present in each tumor. For (D) and (E), values were calculated on rarefied data of 40 samples per tumor type, with 10 iterations. For each iteration, only bacteria that passed all filters in any of the tumor types were included in the analysis. (F) Rarefaction plots for the number of bacterial genera that passed all filters in breast tumor, breast NAT, and breast normal samples. Light color shading represents confidence intervals based on 100 random subsamplings for each number of samples that was analyzed. (G) Fluorescent images from four human breast tumors that were cultured ex vivo with fluorescently labeled D-alanine for 2 hours (blue). Nuclei were stained with DRAQ5 (orange). Scale bars, 10 μ m.

Tumor-adjacent normal breast tissue had an intermediate bacterial load and richness, between those of the breast tumor and normal samples (Fig. 3F and fig. S12). In contrast, we did not find a higher bacterial load in lung and ovarian tumors compared with their tumor-adjacent normal tissues (fig. S12).

To determine whether live bacteria are present in human tumors, we collected fresh breast tumor samples from five women undergoing breast surgery. All tissues were gently dissociated in sterile conditions, plated on 35 types of agar growth media, and incubated in both aerobic and anaerobic conditions, representing a broad span of growth conditions to accommodate a high diversity of bacteria (table S5) (46). In agreement with the positive staining of these tumors for LPS and 16S rRNA FISH (fig. S13), >1000 colonies were grown per tumor from four of the tumors, and 37 colonies were grown from one tumor. In contrast, applying the same steps of tissue dissociation and culturing protocol to five full sets of negative control plates (350 plates) using only phosphate-buffered saline (PBS) yielded only five colonies in total. Whole-genome sequencing of 474 representative colonies from all five tumors demonstrated that they represented 37 different bacterial species, 11 of which (29.7%) are bacteria that were previously detected as hits in our breast tumor cohort (table S5). Fifteen isolated species (40.5%) were detected in our breast tumor cohort but did not pass all filters. For 105 of the colonies, we could not identify the bacteria at the species level (table S5 and materials and methods). Overall, these results show that live bacteria from three main phyla—Proteobacteria, Firmicutes, and Actinobacteria—can be found in breast tumors.

To further validate the presence of live, metabolically active bacteria in human tumors, we cultured slices from four freshly resected human breast tumors ex vivo in the presence of fluorescently labeled D-alanine or dimethyl sulfoxide (DMSO) control. Although D-alanine is used by bacteria to generate peptidoglycan, an essential component of the bacterial cell wall, it is not used by mammalian cells (fig. S14) (47). We detected intracellular labeling in all four tumors, which supports the hypothesis that the tumors harbor live intracellular bacteria (Fig. 3G).

Different tumor types have distinct microbial compositions

Using a single sequencing methodology and platform for the characterization of the microbiome in multiple tumor types enabled us to directly compare the microbiomes of these tumors. Comparison of the beta-diversity between all pairs of samples within a given

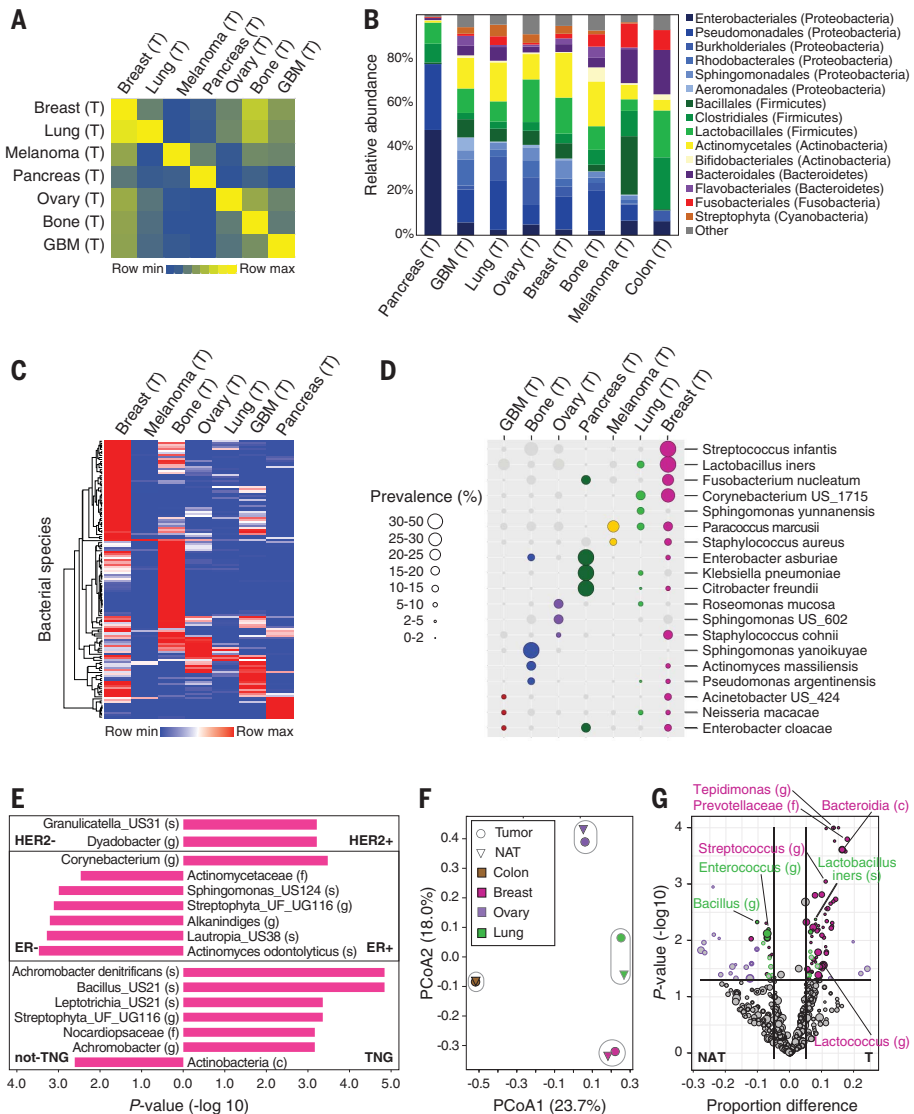


Fig. 4. Different tumor types have distinct microbial compositions. (A) Jaccard similarity indexes were computed on the basis of profiles of bacterial species that passed all filters in tumors ($n = 528$) between all possible pairs of samples. The heatmap presents the average of all indexes between sample pairs from any two cancer types. (B) Distribution of order-level phylotypes across different tumor types. Relative abundances were calculated by summing up the reads of species that passed all filters in the different tumor types and belong to the same order. Orders are colored according to their associated phylum. (C) Unsupervised hierarchical clustering of the prevalence of 137 bacteria species that were hits in one of the tumor types and are present in 10% or more of the samples in at least one of the tumor types. (D) The prevalence of 19 bacteria from (C), displayed across the different tumor types. Only bacteria that are a hit in a given tumor type are represented with colored circles. Circle size indicates the prevalence level in samples. US, unknown species. (E) Bacterial taxa with a significant differential prevalence between different breast tumor subtypes are presented in a bar plot. P values were calculated using a two-sample proportion z test to compare between HER2+ ($n = 61$) and HER2- ($n = 247$), ER+ ($n = 270$) and ER- ($n = 49$), or triple negative (TNG) ($n = 36$) and non-TNG ($n = 284$) breast tumors. The direction of the bars indicates the enrichment direction. All bacteria presented had a false discovery rate (FDR)-corrected Q value <0.25 . US, unknown species; UG, unknown genus; UF, unknown family; (s), species; (g), genus; (f), family; (c), class. (F) Principal coordinate analysis (PCoA) biplot on the Jaccard similarity indexes between bacterial species profiles of the different tissue types. Only bacteria that passed all filters for the specific tissue type were considered. Tumor types and their normal tissue are grouped. (G) Volcano plot demonstrating the differential prevalence of bacteria between tumors (T) and their NAT in breast, lung, and ovary samples. A two-sample proportion z test was used to calculate the P values. Sizes of dots reflect phylotype levels, gradually increasing from species to phylum. Bacteria are colored according to the tumor type (breast, pink; lung, green; and ovary, purple) if they passed significance thresholds (effect size $>5\%$, P value <0.05 , and FDR-corrected Q value <0.25).

tumor type and across different tumor types revealed that the microbiomes of tumors of the same type tend to be more similar to each other than they are to the microbiomes of other tumor types (Fig. 4A and fig. S15). The distribution of order-level phylotypes revealed marked changes between the bacterial composition of the different tumor types (Fig. 4B and fig. S16). We added 22 colorectal tumors from one medical center to our cohort to help relate some of our findings to the known colorectal cancer microbiome (table S2) (11, 12). Consistent with previous reports, bacteria belonging to the Firmicutes and Bacteroidetes phyla were the most abundant species in colorectal tumors (Fig. 4B) (10). In contrast, Proteobacteria dominated the microbiome of pancreatic cancer, similarly to the normal duodenal microbiome makeup (16, 17, 48, 49). This may reflect a retrograde bacterial migration from the duodenum, to which the pancreatic duct opens, as we have previously reported (16). Although species belonging to the Proteobacteria and Firmicutes phyla accounted for most of the detected bacterial sequences in all cancer types, the Proteobacteria to Firmicutes (P/F) ratio appears to vary between tumor types (Fig. 4B). We also detected taxa of the Actinobacteria phylum, including the Corynebacteriaceae and Micrococcaceae families, mostly in nongastrointestinal tumors (Fig. 4B and fig. S16). These observations are in agreement with previous reports describing the microbiome of breast, lung, and ovarian cancer (2, 4, 6, 9, 14, 15, 18).

A tumor-type distinctive microbiome composition was also apparent at the species level. Unsupervised clustering of the most prevalent intratumor bacterial species ($n = 137$ species) demonstrated that many of these species are enriched in certain tumor types (Fig. 4, C and D, and fig. S17). *Fusobacterium nucleatum*, previously reported to be enriched in colorectal tumors, was also a hit in our breast and pancreatic tumor cohorts (fig. S17). We also observed a distinct microbiome across subtypes of the same tumor type. For example, when comparing different subtypes of breast cancer according to their estrogen receptor (ER), progesterone receptor (PR), and HER2 status, we found multiple bacterial taxa whose prevalence was different between the subtypes (Fig. 4E and table S6). Lastly, although the overall microbial composition of the different tumor types was relatively similar to their NAT microbiome (Fig. 4F), we also detected bacteria with a different prevalence in tumors and in their NAT (Fig. 4G and table S7). Consistent with our observation that bacterial load and richness of breast tumors are higher than those in breast NAT (Fig. 3F and fig. S12), we found many bacteria that are significantly enriched in breast tumors compared with their NAT (Fig. 4G).

Metabolic functions encoded by intratumor bacteria are associated with clinical features of certain tumor subtypes

Our results demonstrate that intratumor bacteria span a wide spectrum of the bacterial kingdom. To investigate the functional activities of intratumor bacteria, we used the PICRUSt2 tool (50–52) to map the 16S sequences to the genes and pathways that these bacterial species may harbor (fig. S18 and tables S8 and S9).

Unsupervised clustering analysis of 287 predicted metabolic MetaCyc pathways that showed the greatest variability between the tumor types revealed that certain microbiome metabolic pathways were relatively specific to certain tumor types (Fig. 5A). We found a few tumor type-specific enrichments of bacterial pathways that can degrade metabolites known to be enriched in these same tumor types (table S10). For example, degradation of hydroxyprolines by bacteria (MetaCyc PWY-5159) was enriched in bone

tumors (effect size 14.6%, P value <0.01, proportion test). Bone collagen is a main source of hydroxyproline, and many bone pathologies, like bone tumors, have been shown to result in elevated hydroxyproline levels (53). In the case of lung cancer, MetaCyc pathways responsible for the degradation of chemicals in cigarette smoke, such as toluene, acrylonitrile, and aminobenzoates (TOLUENE-DEG-2-OH-PWY, P344-PWY, and PWY-6077), were significantly enriched in bacteria found in lung tumors compared with other tumor types (effect size 8.4, 8, and 7.2%, P value <0.001 for all, proportion test).

The enrichment for bacteria with the predicted capability to degrade cigarette smoke metabolites in lung tumors may suggest that high levels of these metabolites create a preferred niche for bacteria that can use these metabolites. To confirm this hypothesis, we compared the bacterial functions found in non-small cell lung cancers (NSCLCs) of 100 current

smokers with those in NSCLCs of 43 people who had never smoked (never-smokers). We found that 17 of the 49 MetaCyc pathways that were significantly enriched in tumors of current smokers were pathways that degrade chemicals found in cigarette smoke, such as nicotine, anthranilate, toluene, and phenol (Fig. 5B, blue circles, and table S11). We also found eight MetaCyc pathways related to the biosynthesis of metabolites that can be used by plants—for example, for the biosynthesis of glycine, a key intermediate in plant photorespiration (Fig. 5B, red circles, and table S11). We speculate that some plant-associated bacteria, or their DNA, are present in cigarette tobacco and are thus enriched in the lung tumors of smokers.

To determine which bacteria contribute to the MetaCyc pathways that are enriched in the lung tumors of current smokers, we compared the proportion of all bacterial taxa found in

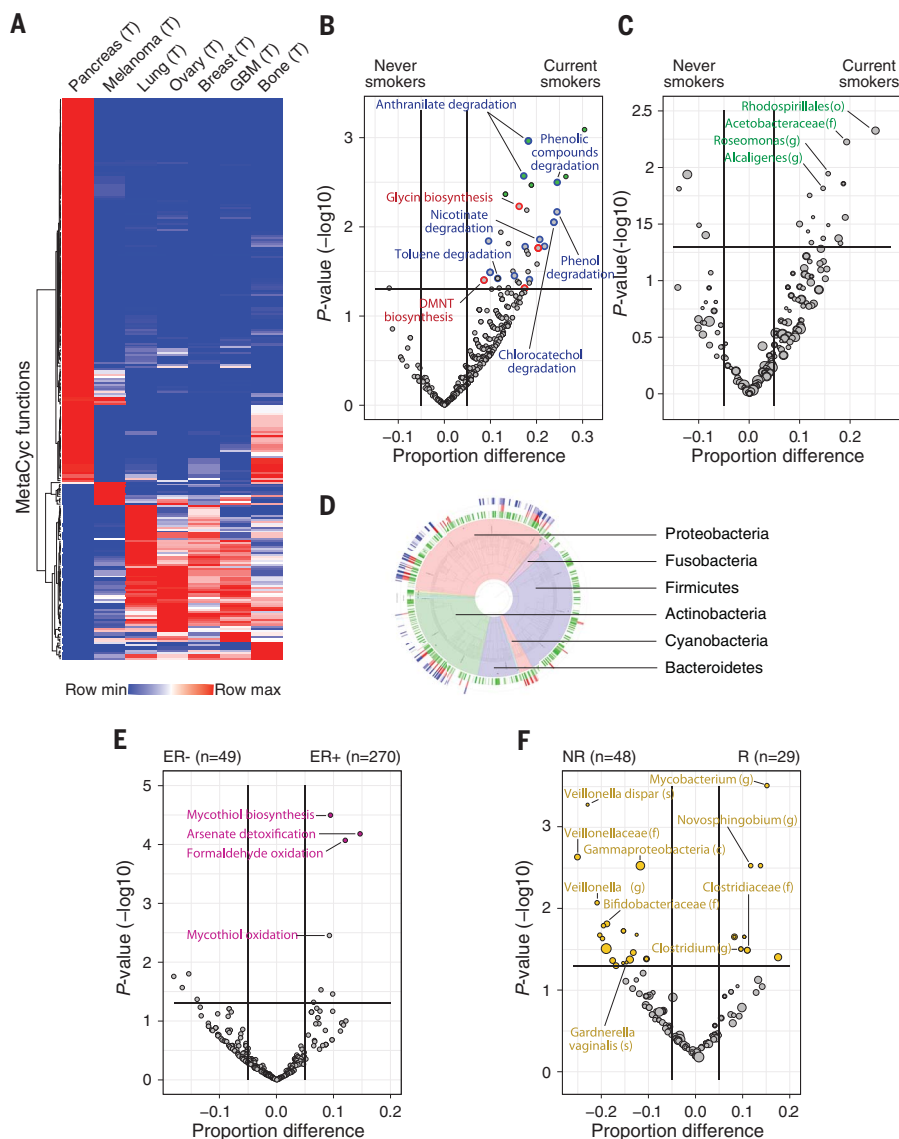


Fig. 5. Predicted bacterial metabolic functions are associated with clinical features. (A) Heatmap demonstrating unsupervised hierarchical clustering of the frequencies of 287 MetaCyc pathways across the different tumor types.

Only pathways that are abundant (frequency >10% in at least one tumor type) and variable (standard deviation divided by the average of frequencies >0.4) were included (table S10). (B and C) Volcano plots demonstrating bacterial MetaCyc pathways (B) and taxa (C) that are enriched in lung tumors from smokers ($n = 100$) versus never-smokers ($n = 43$). Effect size represents the difference in the proportion between the groups. A two-sample proportion z test was used to calculate the P values. Green filled circles indicate pathways with FDR-corrected Q values <0.25. Degrading pathways of smoke chemicals are indicated by blue circles in (B); plant-related metabolic pathways are indicated by red circles in (B). (D) Taxonomy wheel plot of bacterial species that contributed to degradation of cigarette smoke metabolites (blue ring) and to the biosynthesis of plant metabolites functions (red ring) are indicated on the phylogenetic tree, along with all bacteria that are hits in lung tumors (green ring). (E) Volcano plot demonstrating enriched bacterial MetaCyc functions in ER+ versus ER- breast tumors. A two-sample proportion z test was used to calculate the P values. Colored circles indicate pathways with FDR-corrected Q values <0.25. (F) Volcano plot demonstrating the bacterial taxa enriched in melanoma patients who responded (R) to immune checkpoint inhibitors (ICI) versus nonresponders (NR). A binomial test was used to calculate the P values for the enrichment or depletion of bacterial taxa in the responder cohort versus the non-responder cohort. The size of dots reflects phylo-type levels, gradually increasing from species to phylum. Colored circles indicate taxa with FDR-corrected Q values <0.25.

lung tumors of current smokers ($n = 100$) with those in the tumors of never-smokers ($n = 43$). We found that most of the enriched taxa in the lung tumors of smokers belong to the Proteobacteria phylum. However, none of these bacteria reached significance after correction for multiple-hypothesis testing (Fig. 5C and table S12), which indicates that there was no homogeneous population of species conferring this functionality across samples. We reasoned that, although bacterial ecology differs between tumors, there is a shared functional signal related to the specific environment within the lungs of smokers. We were able to demonstrate that a very large number of heterogeneous bacteria contribute to the degradation functions of cigarette smoke metabolites and the biosynthesis of plant metabolites (Fig. 5D). Bacteria expressing these functions are found mainly in the Proteobacteria, Actinobacteria, and Cyanobacteria phyla, and they are depleted from the Firmicutes phylum (Fig. 5D).

We also found selective enrichment of bacterial functions in certain tumor subtypes. For example, multiple MetaCyc pathways were enriched in bacteria from 270 ER+ breast tumors compared with 49 ER– breast tumors (Fig. 5E and table S13). The most significantly enriched pathways in bacteria within ER+ breast tumors were arsenate detoxification and mycothiol biosynthesis. Arsenic is a Group 1 carcinogen that can increase the risk of breast cancer (54) and has been shown to induce expression of the estrogen receptor in human breast cancer (55). Mycothiol is used by bacteria to detoxify reactive oxygen species (56). Because ER+ breast tumors are known to have increased oxidative stress compared with ER– tumors (57), we hypothesize that bacteria with the ability to synthesize mycothiol can better survive in the ER+ tumor microenvironment. We also found enrichment of bacterial functions when comparing breast tumor with NAT samples (table S14). For example, enzymes related to anaerobic respiration were enriched in bacteria from breast cancer versus NAT. Overall, our analysis of MetaCyc pathways suggests a connection between the functions of bacteria present in the tumor and their tumor microenvironment.

Lastly, as our IF staining suggests (Fig. 2), bacteria can be found inside CD45+ immune cells, which indicates that they might influence or reflect the immune state of the tumor microenvironment. To determine whether a specific intratumor microbial signature is correlated with the response to immunotherapy, we compared metastatic melanomas from patients who responded to immune checkpoint inhibitors (ICI) ($n = 29$) with those from patients who did not respond ($n = 48$). Although we did not find significant changes in the load of bacteria between responders and nonresponders to ICI, we did find multiple taxa that

were differentially more ($n = 18$) or less ($n = 28$) abundant in the melanomas of responders compared with nonresponders (Fig. 5F, fig. S19, and table S15). Taxa that were more abundant in tumors of responders included *Clostridium*, whereas *Gardnerella vaginalis* was more abundant in tumors of nonresponders. Notably, this is in line with differential abundances of taxa in the gut microbiome of melanoma patients responding to ICI (23–25).

Discussion

In the present study, we characterized the microbiome of 1526 samples from seven human tumor types. We took multiple measures to minimize and control for contamination (supplementary note) and used our 5R multiplexed bacterial 16S rDNA PCR sequencing technique to gain species-level resolution.

The exploration of multiple tumor types with a single platform allowed us to compare different tumor types and uncover cancer type-specific microbial signatures. This is consistent with a recent publication that demonstrated that reexamination of whole-genome and whole-transcriptome sequencing data from The Cancer Genome Atlas (TCGA) for microbial sequences identified associations between different cancer types and specific microbiota (19). Extending our analysis to the functional level demonstrated that, despite a very large variation in taxa levels, certain tumor environments are enriched for common, relevant bacterial functional traits. This observation is somewhat analogous to the relative stability of the human gut microbiome functions compared with its microbial taxa (58, 59). Using multiple visualization methods and culturomics, we were able to validate the presence of bacteria in the tumors and demonstrate their intracellular localization in both cancer and immune cells.

Our data do not establish whether intratumor bacteria play a causal role in the development of cancer or whether their presence simply reflects infections of established tumors (60, 61). As tumors develop, their disorganized, leaky vasculature may allow circulating bacteria to enter, and the immunosuppressed environment may provide a refuge for them (61, 62). Intratumor bacteria may also arise from the NAT, which can explain the high similarity we found between the tumor microbiome and its NAT microbiome. Whether or not bacteria play a causal role in tumorigenesis, it is of interest to further explore the effects that intratumor bacteria may have on different phenotypes of cancer cells and on the immune system and its interactions with tumor cells. Just as manipulation of the gut microbiome has been shown to affect the response of tumors to immune-checkpoint blockade therapy (23–25, 28), we speculate that manipulation of the tumor microbiome may also affect tumor

immunity and the response to immune therapy. Thus, better understanding of these effects may pave the way for novel treatment options for cancer patients.

REFERENCES AND NOTES

1. C. de Martel et al., *Lancet Oncol.* **13**, 607–615 (2012).
2. C. Xuan et al., *PLOS ONE* **9**, e83744 (2014).
3. K. J. Thompson et al., *PLOS ONE* **12**, e0188873 (2017).
4. S. Banerjee et al., *Front. Microbiol.* **9**, 951 (2018).
5. L. Costantini et al., *Sci. Rep.* **8**, 16893 (2018).
6. K. L. Greathouse et al., *Genome Biol.* **19**, 123 (2018).
7. B. A. Peters et al., *Cancer Epidemiol. Biomark. Prev.* **28**, 731–740 (2019).
8. P. Apostolou et al., *J. Cardiothorac. Surg.* **6**, 137 (2011).
9. S. Banerjee et al., *Oncotarget* **8**, 36225–36245 (2017).
10. Z. Gao, B. Guo, R. Gao, Q. Zhu, H. Qin, *Front. Microbiol.* **6**, 20 (2015).
11. M. Castellari et al., *Genome Res.* **22**, 299–306 (2012).
12. A. D. Kostic et al., *Genome Res.* **22**, 292–298 (2012).
13. K. S. Stefanou et al., *Prostate* **68**, 306–320 (2008).
14. C. Urbaniak et al., *Appl. Environ. Microbiol.* **82**, 5039–5048 (2016).
15. C. Urbaniak et al., *Appl. Environ. Microbiol.* **80**, 3007–3014 (2014).
16. L. T. Geller et al., *Science* **357**, 1156–1160 (2017).
17. S. Pushalkar et al., *Cancer Discov.* **8**, 403–416 (2018).
18. G. Yu et al., *Genome Biol.* **17**, 163 (2016).
19. D. D. Poore et al., *Nature* **579**, 567–574 (2020).
20. B. Goodman, H. Gardner, J. Pathol. **244**, 667–676 (2018).
21. S. H. Wong et al., *Gastroenterology* **153**, 1621–1633.e6 (2017).
22. A. D. Kostic et al., *Cell Host Microbe* **14**, 207–215 (2013).
23. V. Gopalakrishnan et al., *Science* **359**, 97–103 (2018).
24. V. Matson et al., *Science* **359**, 104–108 (2018).
25. B. Routy et al., *Science* **359**, 91–97 (2018).
26. S. Viaud et al., *Science* **342**, 971–976 (2013).
27. N. Iida et al., *Science* **342**, 967–970 (2013).
28. B. A. Helmink, M. A. W. Khan, A. Hermann, V. Gopalakrishnan, J. A. Wargo, *Nat. Med.* **25**, 377–388 (2019).
29. M. R. Rutkowski et al., *Cancer Cell* **27**, 27–40 (2015).
30. S. J. Salter et al., *BMC Biol.* **12**, 87 (2014).
31. R. Eisenhofer et al., *Trends Microbiol.* **27**, 105–117 (2019).
32. N. M. Davis, D. M. Proctor, S. P. Holmes, D. A. Relman, B. J. Callahan, *Microbiome* **6**, 226 (2018).
33. M. C. de Goffau et al., *Nature* **572**, 329–334 (2019).
34. M. L. Sogin et al., *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12115–12120 (2006).
35. C. R. H. Raetz, C. Whitfield, *Annu. Rev. Biochem.* **71**, 635–700 (2002).
36. W. Fischer, in *New Comprehensive Biochemistry*, vol. 27, J.-M. Ghuysen, R. Hakenbeck, Eds. (Elsevier, 1994), pp. 199–215.
37. R. I. Amann et al., *Appl. Environ. Microbiol.* **56**, 1919–1925 (1990).
38. C. Forestier, E. Moreno, J. Pizarro-Cerda, J. P. Gorvel, *J. Immunol.* **162**, 6784–6791 (1999).
39. K. T. Tokuyasu, *J. Microsc.* **143**, 139–149 (1986).
40. A. Abada, S. Levin-Zaidman, Z. Porat, T. Dadosh, Z. Elazar, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 12749–12754 (2017).
41. E. Klieneberger-Nobel, *J. Gen. Microbiol.* **3**, 434–443 (1949).
42. J. Errington, *Open Biol.* **3**, 120143 (2013).
43. J. Errington, *Biochem. Soc. Trans.* **45**, 287–295 (2017).
44. G. Fuks et al., *Microbiome* **6**, 17 (2018).
45. T. Z. DeSantis et al., *Appl. Environ. Microbiol.* **72**, 5069–5072 (2006).
46. J. T. Lau et al., *Genome Med.* **8**, 72 (2016).
47. M. S. Siegrist et al., *ACS Chem. Biol.* **8**, 500–505 (2013).
48. G. Ou et al., *Am. J. Gastroenterol.* **104**, 3058–3067 (2009).
49. E. Nistal et al., *Inflamm. Bowel Dis.* **18**, 649–656 (2012).
50. G. M. Douglas et al., PICRUSt2: An improved and extensible approach for metagenome inference, *bioRxiv* 672295 [Preprint]. 15 June 2019. www.biorxiv.org/content/10.1101/672295v1.
51. Y. Ye, T. G. Doak, *PLOS Comput. Biol.* **5**, e1000465 (2009).
52. S. Louca, M. Doebeli, *Bioinformatics* **34**, 1053–1055 (2018).
53. M. Nakagawa, Y. Sugiura, T. Oshima, G. Kajino, H. Hirako, *Nagoya J. Med. Sci.* **29**, 345–367 (1967).
54. N. Khanjani, A.-B. Jafarnejad, L. Tavakkoli, *Rev. Environ. Health* **32**, 267–277 (2017).
55. J. Du et al., *PLOS ONE* **7**, e35957 (2012).

56. A. M. Reyes *et al.*, *Antioxid. Redox Signal.* **28**, 487–504 (2018).
57. P. Karihtala, S. Kauppila, Y. Soini, Arja-Jukkola-Vuorinen, *BMC Cancer* **11**, 262 (2011).
58. Human Microbiome Project Consortium, *Nature* **486**, 207–214 (2012).
59. A. Eng, E. Borenstein, *Microbiome* **6**, 45 (2018).
60. J. L. Pope, S. Tomkovich, Y. Yang, C. Jobin, *Transl. Res.* **179**, 139–154 (2017).
61. J. Cummins, M. Tangney, *Infect. Agent. Cancer* **8**, 11 (2013).
62. C. K. Baban, M. Cronin, D. O'Hanlon, G. C. O'Sullivan, M. Tangney, *Bioeng. Bugs* **1**, 385–394 (2010).
63. N. Shental, NoamShental/5R: First release, April 2020, version v1.0, Zenodo (2020); <http://doi.org/10.5281/zenodo.3740525>.
64. I. Liviyatan, iliviyatan/TMB: TMB for Science publication, version science_pub, Zenodo (2020); <http://doi.org/10.5281/zenodo.3740536>.

ACKNOWLEDGMENTS

We thank all members of the Straussman laboratory as well as E. Elinav, D. Douek, U. Gophna, and R. Kolter for fruitful discussions. **Funding:** R.St. is funded by the Israel Science Foundation (grant no. 2044/17), the Binational Science Foundation (grant no. 2013332), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 818086), the Fabrikant-Morse Families Research Fund for Humanity, the Chantal d'Adesky Scheinberg Research Fund, the Moross Integrated Cancer Center, the Rising Tide Foundation, the International Collaboration Grant from the Jacki and Bruce Barron Cancer Research Scholars' Program, and a partnership of the Israel Cancer Research Fund and City of Hope (COH), as supported by The Harvey L. Miller Family Foundation. R.St. is the incumbent of the Roel C. Buck Career Development Chair. N.S. is funded by the Ministry of Science, Technology and Space, Israel (grant no. 3-11174).

Research reported in this publication included work performed in the COH Pathology Research Services Core supported by the National Cancer Institute of the National Institutes of Health under award no. P30CA033572. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. A.P.C. is supported by the Cancer Prevention and Research Institute of Texas Research Training Program (RP170067). **Author contributions:** D.N., I.L., G.F., N.S., and R.St. conceived and initiated the project and performed most of the data analysis. I.L., G.F., and O.G. wrote data analysis code. D.N., N.G., and Y.Z. performed staining experiments. D.N., L.T.G., and A.R.-M. generated sequencing data. R.W. and N.G. performed the culturomics experiments. A.W., M.L.-P., Y.B., L.T.G., A.R.-M., E.G., A.M., and E.S. contributed to the culturomics experiment. G.M.D., M.G.I.L., and I.L. were responsible for transforming bacterial species into functions using PICRUSt 2. I.L. and T.A. generated GraPhlAn analysis. T.Da. and S.L.-Z. performed CLEM experiments. R.St., D.N., I.L., and N.S. wrote the initial manuscript. G.F., N.G., and Y.Z. helped with preparations of the manuscript. J.S., I.K., V.G., S.A., G.P., M.R., Z.A.C., R.A., A.P.C., M.A.W.K., G.O., K.B.-S., S.H., T.S., E.A.R., C.U.B., A.R., R.Sh., A.A., T.Do., R.K., Z.R.C., E.N.G.-Y., E.Y.-S., H.S., N.B., A.B.-N., B.K., A.N., T.G., M.D., K.L., J.B., S.Y.-K., I.B., D.S.P., R.W., J.A.W., and D.J.R. provided patient samples. J.S. and G.M. helped with pathological analysis. R.St. supervised the project. **Competing interests:** R.St. received a grant from Merck EMD Serono and is a paid adviser to Biomica and BiomX. R.St., N.S., D.N., I.L., and G.F. are inventors on a U.S. provisional patent application (63/005,540) submitted by Yeda Research and Development, the Weizmann Institute of Science. J.A.W. and V.G. are inventors on a U.S. patent application (PCT/US17/53717) submitted by the University of Texas MD Anderson Cancer Center that covers methods to enhance immune checkpoint blockade responses by modulating the microbiome. J.A.W. reports compensation for speaker's bureau and honoraria from Imedex, Dava Oncology, Omniprex, Illumina,

Gilead, PeerView, Physician Education Resource, MedImmune, and Bristol-Myers Squibb. J.A.W. serves as a consultant and advisory board member for Roche/Genentech, Novartis, AstraZeneca, GlaxoSmithKline, Bristol-Myers Squibb, Merck, Biothera Pharmaceuticals, and Microbiome DX. J.A.W. also receives research support from GlaxoSmithKline, Roche/Genentech, Bristol-Myers Squibb, and Novartis. **Data and materials availability:** Breast and colon samples from Sheba; melanoma samples from the Netherlands Cancer Institute–Antoni van Leeuwenhoekziekenhuis (NKI-AVL); and lung, breast, ovary, and GBM samples from the Israeli Biorepository Network for Research (MIDGAM) are available from R.St. under material transfer agreements with the Weizmann Institute. All processed data are available in the manuscript or the supplementary materials. The software package for applying SMURF using the 5R primers is available at Zenodo (63). The implementation code for the 16S sequencing analysis pipeline with filters can be found at Zenodo (64). Bacterial 16S sequencing data are available from the National Center for Biotechnology Information BioProject database under accession no. PRJNA624822.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6494/973/suppl/DC1
 Supplementary Note
 Materials and Methods
 Figs. S1 to S19
 Tables S1 to S15
 References (65–76)
 MDAR Reproducibility Checklist

[View/request a protocol for this paper from Bio-protocol.](#)

29 July 2019; resubmitted 22 January 2020
 Accepted 14 April 2020
 10.1126/science.aay9189

PEPTIDE SYNTHESIS

Synthesis of proteins by automated flow chemistry

N. Hartrampf^{1*}, A. Saebi^{1†}, M. Poskus^{1‡}, Z. P. Gates^{1‡}, A. J. Callahan¹, A. E. Cowfer¹, S. Hanna¹, S. Antilla^{1,2}, C. K. Schissel¹, A. J. Quartararo¹, X. Ye¹, A. J. Mijalis^{1§}, M. D. Simon^{1¶}, A. Loas¹, S. Liu^{1#}, C. Jessen³, T. E. Nielsen³, B. L. Pentelute^{1**}

Ribosomes can produce proteins in minutes and are largely constrained to proteinogenic amino acids. Here, we report highly efficient chemistry matched with an automated fast-flow instrument for the direct manufacturing of peptide chains up to 164 amino acids long over 327 consecutive reactions. The machine is rapid: Peptide chain elongation is complete in hours. We demonstrate the utility of this approach by the chemical synthesis of nine different protein chains that represent enzymes, structural units, and regulatory factors. After purification and folding, the synthetic materials display biophysical and enzymatic properties comparable to the biologically expressed proteins. High-fidelity automated flow chemistry is an alternative for producing single-domain proteins without the ribosome.

Mechanical pumps, valves, solid supports, and computers have transformed the way we perform chemical reactions. Continuous, multistep flow technology has enabled routine access to small molecules ranging from pharmaceutical ingredients to natural products and bulk commodities (1). Advantages of flow synthesis over batch methods are in-line spectroscopic monitoring, efficient mixing, and precise control over the reaction parameters (2). Translating these capabilities to the total chemical synthesis of peptides and proteins will provide rapid access to an expanded chemical space.

Protein production is an essential part of research in academia and industry and can be accomplished by biological methods or chemical synthesis. Most proteins are obtained by biological expression, a process that largely limits their chemical composition to the canonical proteinogenic amino acids (3). Advances in genetic code expansion have allowed for the incorporation of up to two unnatural amino acids in the structures of native proteins (4). By contrast, chemical synthesis offers unmatched flexibility when incorporation of multiple unnatural amino acids, posttranslational modifications, or artificial backbones is desired (3). Synthetic proteins have become accessible with a combination

of solid-phase and ligation methodologies. Yet, total chemical synthesis of proteins remains highly labor intensive.

Solid-phase peptide synthesis (SPPS) is the foundation of chemical peptide and protein production (5). Despite decades of optimization, peptides longer than 50 amino acids are difficult to synthesize with standard SPPS instrumentation, owing in large part to generation of by-products from deletion, truncation, and aggregation of the growing chains (6, 7). It was not until the development of native chemical ligation (NCL) that chemical synthesis of protein chains became practical (8). Despite the efforts dedicated to improving NCL techniques (9), a major bottleneck resides in the absence of a routine, widely applicable protocol to access the requisite peptide fragments (10, 11). We set out to address this problem by developing a reliable method to synthesize long peptides and protein chains using flow chemistry.

Flow-based SPPS is gaining momentum owing to its advantageous features—for example, control over physical parameters and greatly reduced formation of side products (12–14). Studies carried out as early as 1970 found that automation and high fidelity of peptide synthesis could be achieved by containing the solid support in a reactor and operating it as a fixed bed (15, 16). Instead of complex systems for liquid handling to dispense reagents and wash the resin, high-performance liquid chromatography (HPLC) pumps were used to continuously deliver reagents, establishing the principles of peptide synthesis in flow. Inspired by this early work, over the past 5 years we developed rapid, automated fast-flow peptide synthesis (AFPS) instrumentation that incorporates amino acid residues in as little as 40 s at temperatures up to 90°C (17–19).

Even though prior work by us and others on flow-based SPPS considerably reduced the total synthesis time, the potential of flow chemistry to enable synthesis of peptide chains in

the range of single-domain proteins has not been fully realized (17–23). We set out to optimize our AFPS technology to meet this challenge (19). We report here a routine protocol that allows for stepwise chemical total synthesis of peptide chains exceeding 50 amino acids in length, with a cycle time of ~2.5 min per amino acid (Fig. 1A). The optimized protocol was built on a collection of analytical data acquired with an AFPS system and delivers products with high fidelity and of high chiral purity. Using this protocol, single-domain protein chains ranging from barstar (90 amino acids) to sortase A_{59–206} (sortase A*, 164 amino acids) were synthesized in 3.5 to 6.5 hours. To demonstrate production of functional proteins, these sequences were folded, and their biophysical properties and enzymatic activities were determined. The time scale of chemical protein synthesis is on par with that of recombinant expression and therefore offers a practical alternative to biochemical methods while opening up the chemical space beyond canonical amino acids.

Rapid screening of reaction variables for AFPS protocol development

We chose to first optimize coupling efficiency and later investigate possible side reactions induced by the optimized coupling conditions. On a benchmark AFPS instrument previously developed in our laboratory (17, 19), reagents are mixed, heated, and delivered onto a pre-tempered solid support using three HPLC pumps. In-line ultraviolet-visible (UV-vis) detection of the reactor eluent is used to monitor removal of the N-terminal protecting group after each coupling cycle. Indirectly, this information reports on the efficiency of the preceding coupling step.

We first optimized general parameters, including flow rate, reaction solvent, reagent concentration, temperature, and coupling agents (Fig. 1B and tables S1 to S7). Modifications to our original AFPS protocol included increasing reagent concentrations to 0.4 M (24), the use of amine-free *N,N*-dimethylformamide (DMF), and an increase in temperature to 85° to 90°C for reagent activation and coupling. The performance of different activators for the coupling step was also investigated, identifying the azabenzotriazol-reagents PyAOP [(7-azabenzotriazol-1-yloxy)tripyrrolidino-phosphonium hexafluorophosphate] and HATU (hexafluorophosphate azabenzotriazole tetramethyl uranium) as optimal.

Automated collection and analysis of data combined with the synthesis parameters allowed for optimization of residue-specific coupling conditions. By comparing data on amino acid deprotections, we were able to gain information on coupling efficiency for all canonical amino acids and generated a general amino acid-specific recipe (tables S8 and S9).

¹Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. ²Department of Materials Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA.

³Novo Nordisk A/S, Novo Nordisk Park, DK-2760 Måløv, Denmark.

*Present address: Department of Chemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

†These authors contributed equally to this work.

‡Present address: p53Lab, Agency for Science, Technology and Research (A*STAR), 8A Biomedical Grove, Singapore 138648, Singapore.

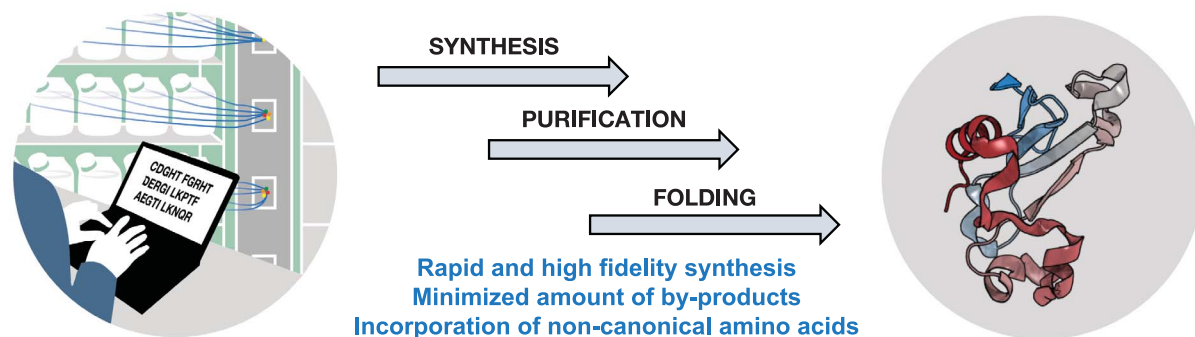
§Present address: Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA.

¶Present address: Ginkgo Bioworks, 27 Drydock Avenue, Boston, MA 02210, USA.

#Present address: Department of Chemistry, East China Normal University, 3663 North Zhongshan Road, Shanghai 200062, China.

**Corresponding author. Email: blp@mit.edu

A Fully Automated Synthesis of Functional Synthetic Proteins



B HPLC traces of GLP-1 before and after optimization of synthesis conditions

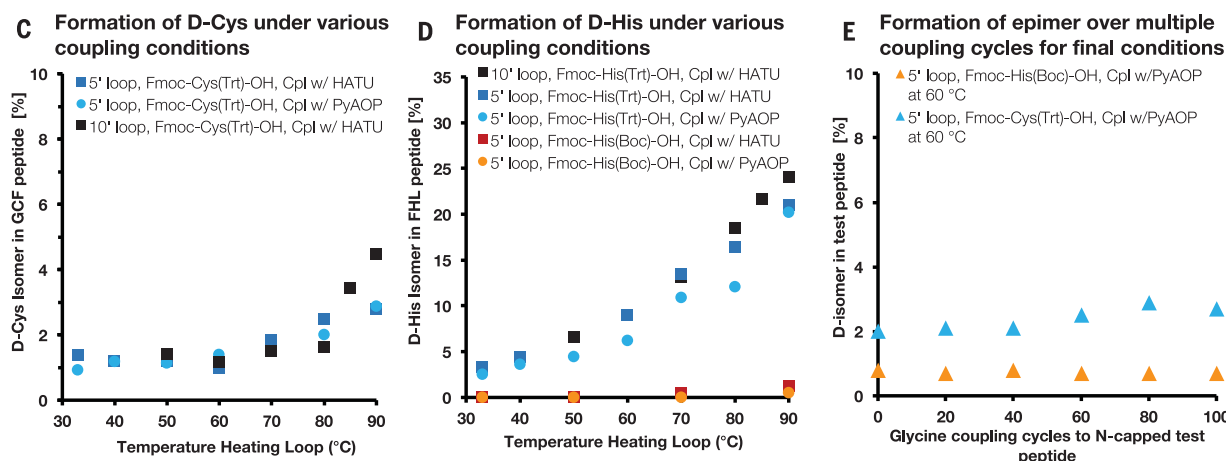
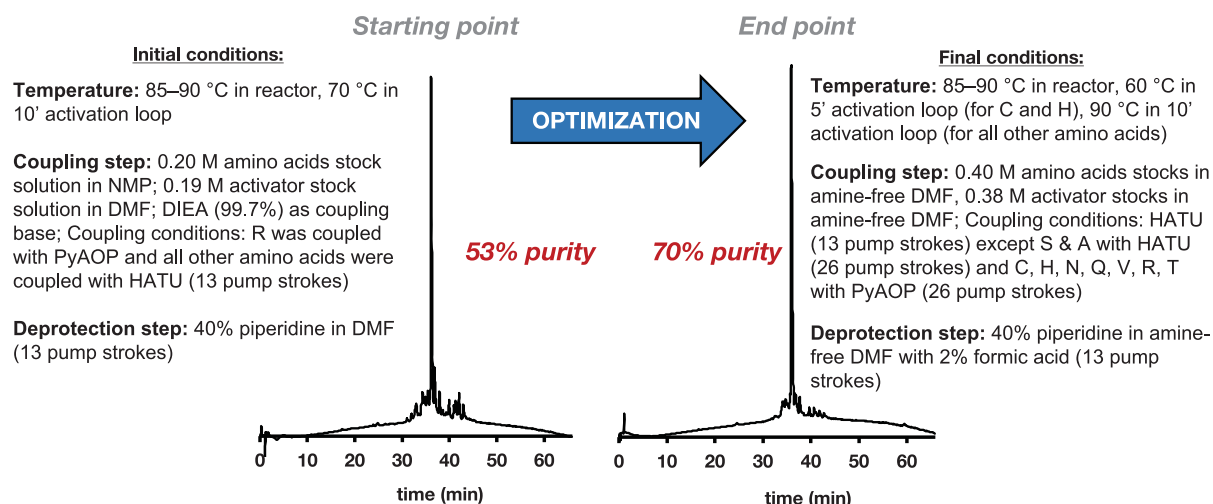


Fig. 1. Optimized conditions for automated fast-flow solid-phase peptide synthesis enable high-fidelity production of long amino acid sequences.

(A) Fully automated chemical flow synthesis yields peptide chains, which, after purification and folding, give functional proteins. Protein Data Bank (PDB) 1BR5 (barnase) (57) was used. (B) Synthesis of GLP-1 using starting conditions and optimized conditions. The concentrations listed refer to stock solutions. NMP, *N*-methylpyrrolidone; DIEA, diisopropylethylamine. (C) Quantification of cysteine epimerization as a function of activation temperature, heating time (5' loop and 10' loop), and activator in a GCF test peptide. Isomer was quantified from extracted ion chromatograms on LC-MS by comparison to reference peptides.

(D) Quantification of histidine epimerization as a function of activation temperature, heating time (5' loop and 10' loop), and activator in a FHL test peptide. The D-isomer was quantified by analytical HPLC by comparison to reference peptides. (E) Quantification of epimerization over multiple coupling cycles. GCF and FHL were synthesized under optimized conditions, and the N terminus was manually capped with a Boc-protecting group. One-hundred glycine couplings were executed, and a sample was taken out for analysis every 20 amino acid couplings. Cpl w/, coupled with. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

Analytical comparison of the products obtained for glucagon-like peptide-1 (GLP-1) is illustrative of the improvement in crude peptide quality achieved with the optimized synthesis conditions (Fig. 1B).

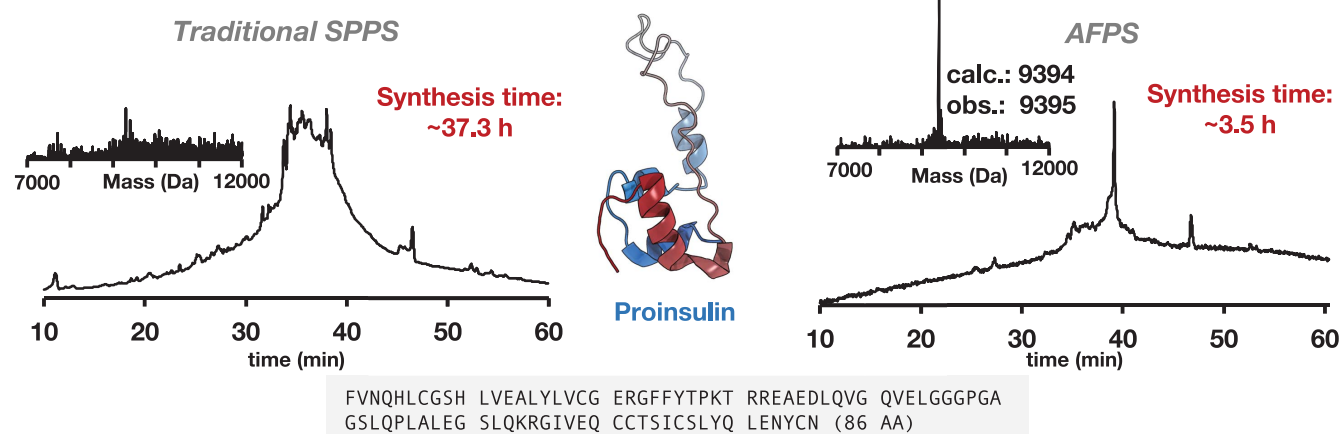
We aimed to suppress aspartimide formation, a major side reaction in SPPS and AFPS. Because increased temperature leads to more aspartimide formation, various deprotection bases, additives, and aspartic acid protecting groups were screened to minimize this unwanted side reaction (25, 26). We found that milder deprotection bases [i.e., piperazine and 1-hydroxybenzotriazole (HOBt) with piperidine] and bulky aspartic acid protecting groups [i.e., *O*-3-methylpent-3-yl (OMpe)] decreased the level of aspartimide formation (fig. S3 and table S10). The most effective strategies, how-

ever, were the addition of formic acid as a piperidine additive and backbone protection with dimethoxybenzyl glycine. Formic acid (2% stock solution in 40:60 v/v piperidine:DMF) was therefore used as an additive for deprotection, and backbone protection was applied for collagen and fibroblast growth factor 1 (FGF1) syntheses.

We confirmed retention of chirality for amino acids at high risk of epimerization (i.e., cysteine and histidine) in a final optimization step (figs. S4 to S9) (27). The influence of temperature, time, and activating agent, as well as different side-chain protecting groups were screened (Fig. 1, C and D) (17). For both amino acids, epimerization increases with activation time and temperature. The choice of protecting group proved to be critical for histidine. Ulti-

mately, activation of Fmoc-Cys(Trt)-OH and Fmoc-His(Boc)-OH with PyAOP with a shorter time at 60°C resulted in <2% D-epimer formation (Fmoc, fluorenylmethyloxycarbonyl; Trt, trityl; Boc, *tert*-butoyloxycarbonyl). Next, we determined that the amount of epimerization under these optimized conditions does not increase over multiple coupling cycles (Fig. 1E). The amount of D-isomer did not change over 100 amino acid couplings executed after manual capping of the N-terminus, indicating that epimerization of cysteine and histidine only occurs during the activation step. Implementation of these conditions allowed us to solidify the general AFPS protocol, which was then applied to the production of sequences exceeding 50 amino acids [table S11 and supplementary materials (SM) section 3.10].

A Traditional SPPS vs. AFPS, analytical data for crude proinsulin



B Traditional SPPS vs. AFPS, analytical data for crude HIV-1 protease

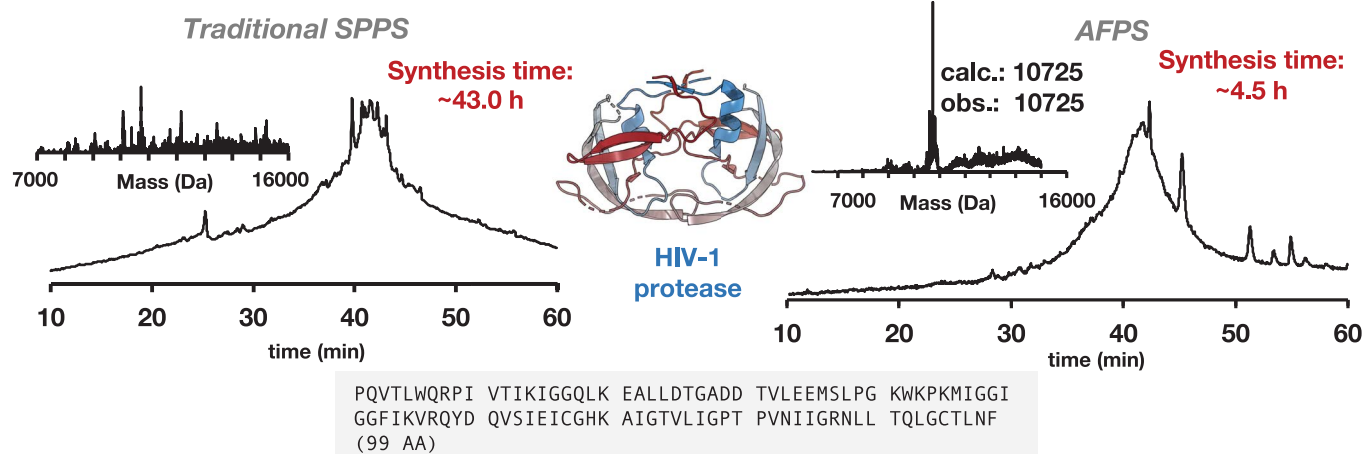


Fig. 2. Synthesis of proinsulin and HIV-1 protease demonstrates the advantage of AFPS over traditional SPPS methods. (A and B) Analytical HPLC data of the crude proinsulin (A) and HIV-1 protease (B) are presented as the main chromatographic traces with absorbance detection at 214 nm (additional details in the SM). Deconvoluted masses are displayed in the insets. Analytical data for the synthesis of crude protein chain using SPPS on a commercially available synthesizer at 70°C with total cycle times of 26 min per amino acid and

40 equivalents of amino acid for each coupling are displayed on the left; analytical data for the synthesis of crude protein chain using AFPS at 90°C with 60 equivalents of amino acid for each coupling are displayed on the right. PDB 2KQP (proinsulin) (58) and 2JE4 (HIV-1 protease dimer with inhibitor, not identical to the sequence synthesized) (30) were used. Crude material of higher quality was obtained in the case of HIV-1 protease by substitution of Cys and Met residues, as described in SM section 5. AA, amino acids.

Optimized AFPS outperforms traditional synthesis methods

We investigated if our optimized AFPS conditions could facilitate the synthesis of longer sequences using proinsulin (86 amino acids) and human immunodeficiency virus-1 (HIV-1) protease (99 amino acids) as test sequences. The total synthesis of human proinsulin was previously reported using NCL of three peptide fragments individually prepared by SPPS (28). HIV-1 protease was previously prepared using stepwise and chemical ligation routes under Boc-SPPS conditions (29, 30). Using our standard AFPS protocol, the syntheses of proinsulin and HIV-1 protease were completed in 3.5 and 4.5 hours, respectively. HPLC purification yielded 2.2 mg (1%) of purified proinsulin and 5.3 mg (1%) of purified HIV-1 protease.

A comparison between AFPS and standard batch SPPS syntheses on commercially available synthesizers at room temperature, 70°C, and 90°C indicated substantially improved synthetic outcome for the optimized AFPS protocol (Fig. 2 and SM section 4). On each instrument, machine-specific, optimized conditions were used to achieve the best synthesis outcome. For HIV-1 protease and proinsulin, AFPS yielded the desired product as the major species along with minor by-products of similar weight, as determined by analytical HPLC and liquid chromatography–mass spectrometry (LC-MS). By contrast, synthesis on commercially available peptide synthesizers took approximately five times longer and resulted in a complex compound mixture. AFPS therefore offers a substantial improvement when directly compared with traditional SPPS methods, both with respect to time and performance.

Optimized AFPS enables routine access to single-domain protein chains

To demonstrate general applicability of our AFPS protocol, the synthesis of additional protein chains ranging from ~70 to ~170 amino acids was performed (Fig. 3A and SM section 5). These sequences were chosen to enable comparison with literature data. We chose not only historically relevant targets for drug discovery, such as HIV-1 protease and murine double minute 2 (MDM2) (31, 32), but also proteins that serve as therapeutics themselves, such as FGFT and proinsulin (33, 34). Barstar, barnase, lysozyme, MDM2, and sortase A* allowed for a direct comparison of recombinant and synthetic proteins. The ability of AFPS technology to rapidly and simultaneously incorporate noncanonical amino acids in greater number and of greater diversity than biological methods was tested by synthesizing derivatives of barnase and HIV-1 protease containing site-directed mutations. In the case of barnase, we incorporated *p*-bromophenylalanine at a site previously investigated for mutational toler-

ance (35). Then, we produced synthetic HIV-1 protease in which two methionine and one cysteine residues were replaced as previously described with norleucine and aminobutyric acid, respectively (Fig. 3B), to avoid potential oxidation side products and increase synthetic efficiency (30). All sequences were successfully synthesized in 3.5 to 6.5 hours.

The desired protein was the main product in every synthesis, and HPLC purification yielded milligram quantities of product. Isolated yields after HPLC purification ranged from 2.2 to 19.0 mg (1 to 5%), a sufficient amount of material for folding and evaluation of tertiary structure and biological function (Fig. 3B and SM section 5). In conclusion, optimized AFPS allows for the routine stepwise chemical synthesis of peptide chains of up to ~170 amino acids and therefore substantially decreases time and labor associated with the chemical production of single-domain proteins.

The structure and function of folded synthetic proteins are comparable to recombinant samples

Determining the purity of long synthetic peptides is challenging because of difficulties associated with identification and quantification of by-products by standard analytical techniques. In a physiological environment, the native folded structure of a globular protein, which gives rise to its distinctive biological activity, is determined by its amino acid sequence (36). As a consequence, the tertiary structure of a protein can be used as a measure of the chemical integrity of the primary amino acid sequence (37).

We folded and purified selected synthetic proteins by size exclusion chromatography and ion exchange chromatography and characterized their tertiary structure with biophysical and functional assays, alongside recombinant protein standards. Our goal was to demonstrate the fidelity of our AFPS protocol in delivering synthetic proteins of defined covalent structure and high chiral integrity. To this aim, we thoroughly characterized barnase and further investigated barstar, sortase A*, MDM2, and HIV-1 protease. Folding of the synthetic proteins was case-specific and was achieved either by following a literature protocol or by screening various conditions.

Chemical denaturation is diagnostic for assessing structural integrity and stability of synthetic proteins. The globular protein barnase, a bacterial ribonuclease (RNase) isolated from *Bacillus amyloliquefaciens*, is a model system to investigate protein folding, denaturation, and binding to its inhibitor protein barstar (Fig. 4A) (38, 39). The primary structures of synthetic and recombinant barnase were indistinguishable by LC-MS and HPLC methods (Fig. 4B). We used a chemical denaturation fluorometric assay as a readout for the in-

tegrity of the tertiary structure (Fig. 4C). In this assay, tryptophan fluorescence was used to monitor the folding equilibrium, as the concentration of urea was varied. Synthetic barnase exhibited a transition midpoint (the concentration at which half of the sample is unfolded, $[D]_{50\%}$) that compared well to both the authentic recombinant sample and literature value $\{[D]_{50\%, \text{synthetic}} = 4.68 \pm 0.06 \text{ M}; [D]_{50\%, \text{recombinant}} = 4.63 \pm 0.04 \text{ M (mean} \pm \text{SE)}; [D]_{50\%, \text{literature}} = 4.57 \text{ M}\}$ (39). More importantly, the m values obtained in the experiment, which describe the slope of the unfolding transition and are a sensitive measure of structural homogeneity, were similar $[m_{\text{synthetic}} = 1.82 \pm 0.25 \text{ kcal mol}^{-1} \text{ M}^{-1}; m_{\text{recombinant}} = 1.88 \pm 0.21 \text{ kcal mol}^{-1} \text{ M}^{-1} \text{ (mean} \pm \text{SE)}; m_{\text{literature}} = 2.06 \text{ kcal mol}^{-1} \text{ M}^{-1}]$ (39). If the synthetic protein were microheterogeneous (e.g., contained a distribution of isomers or deletion coproducts), then the apparent m value may be altered owing to the distribution of $[D]_{50\%}$ values represented within the mixture. Therefore, because the synthetic sample exhibited an m value within the error of the recombinant sample, we concluded that microheterogeneity was negligible.

Enzymatic assays show comparable activity of synthetic proteins obtained by AFPS and their recombinant equivalents. Enzymatic catalysis is sensitive to minor changes in the enzyme's tertiary structure, for which even single point mutations can have a major impact (40, 41). We evaluated the native activity of three synthetic variants of well-studied enzymes: barnase, HIV-1 protease, and sortase A*. Barnase catalyzes hydrolysis at diribonucleotide GpN sites. Its specific activity can be measured by monitoring hydrolysis of a DNA-RNA hybrid containing a Förster resonance energy transfer fluorophore pair (42). The enzymatic efficiency of synthetic barnase was $k_{\text{cat}}/K_M = (7.6 \pm 0.2) \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ (mean $\pm$ SE), which is comparable to that of recombinant barnase $[k_{\text{cat}}/K_M = (9.0 \pm 0.3) \times 10^6 \text{ M}^{-1} \text{ s}^{-1} \text{ (mean} \pm \text{SE)}]$ determined using the same assay (Fig. 4D).

The primary structure of HIV-1 protease was confirmed by LC-MS and HPLC methods (Fig. 5B). HIV-1 protease hydrolyzes the peptides of HIV, and using a fluorogenic peptide allows for quantification of its proteolytic activity (43). Synthetic HIV-1 protease displays a Michaelis constant of $K_M = 20.9 \pm 1.0 \mu\text{M}$ (mean $\pm$ SE) and a turnover number of $k_{\text{cat}} = 29.6 \pm 4.1 \text{ s}^{-1}$ (mean $\pm$ SE), close to literature values published for a similar synthetic sample obtained by SPPS (Fig. 5C) (30). Incubation of the synthetic protease with a model substrate peptide results in wild type-like specificity with exclusive cleavage at a single Phe-Pro site (Fig. 5D) (29).

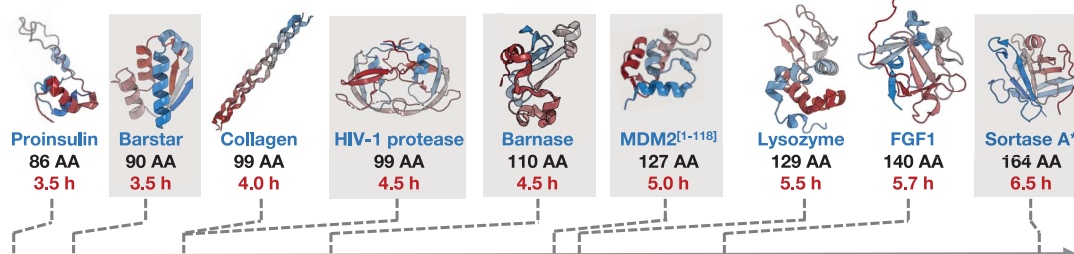
Sortase A_{59–206} is a transpeptidase produced by Gram-positive bacteria that catalyzes a cell wall sorting reaction at a threonine-glycine

bond in the LPXTG motif (Leu-Pro-X-Thr-Gly, where X is any amino acid) (44). We synthesized the 164-amino acid-long sortase A* variant (P94S/D160N/K196T; P, Pro; S, Ser; D, Asp; N, Asn; K, Lys; T, Thr) to allow for direct

comparison to a recombinant standard (45, 46). At a concentration of 0.01 mg/ml, synthetic sortase A* led to 47% product formation by LC-MS within 24 hours (starting from 0.2 mg/ml GGGGGLY and AQLPETGEE as test sub-

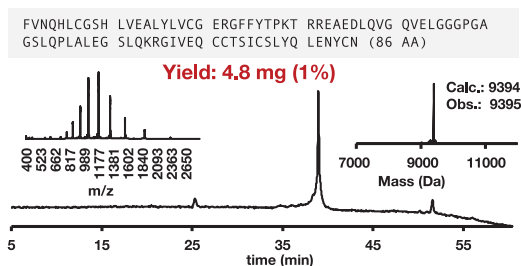
strates; G, Gly; L, Leu; Y, Tyr; A, Ala; Q, Gln; E, Glu) (fig. S23). This conversion value is comparable to that determined for the recombinant protein (50% product formation within 24 hours). Enzymatic activity assays of synthetic

A Single domain proteins synthesized using AFPS

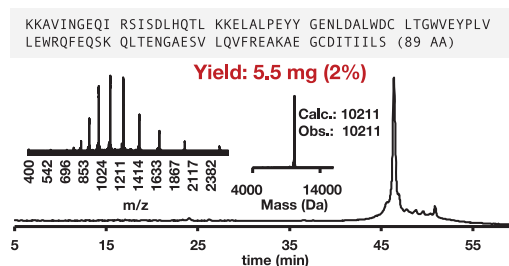


B Analytical data

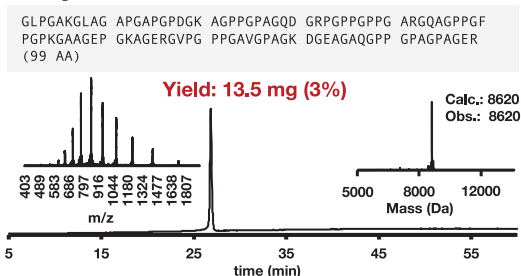
Proinsulin



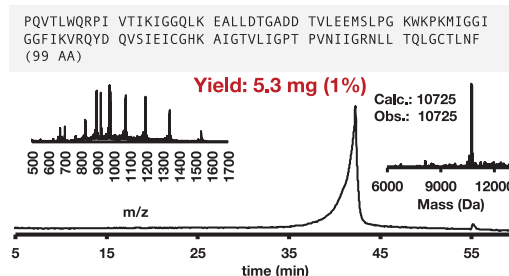
Barstar



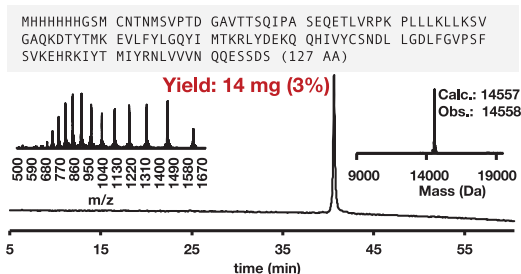
Collagen



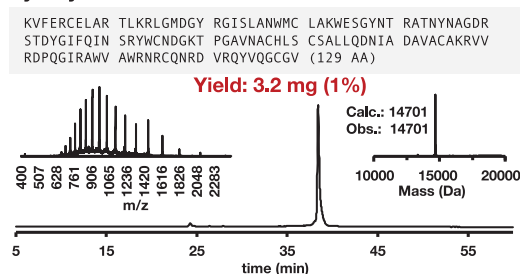
HIV-1 protease



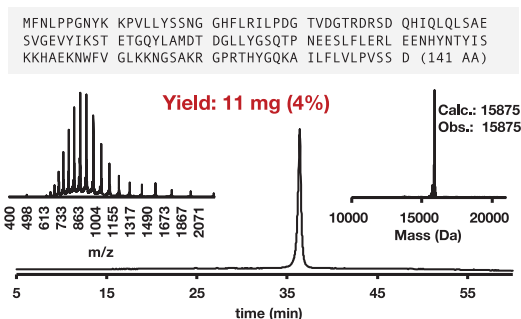
MDM2^[1-118]



Lysozyme



FGF1



Sortase A*

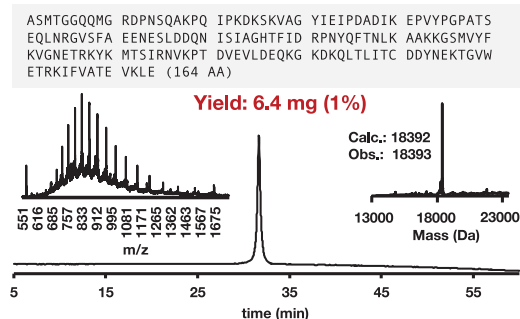


Fig. 3. AFPS enables high-fidelity production of long amino acid sequences in hours. (A) Sequences produced using an AFPS instrument. Sequences high-lighted in gray were folded and purified, and their structure and biological activity were evaluated. All sequences were synthesized using the same standard recipe. PDB 1AY7 (barstar) (59), 2KQP (proinsulin) (58), 1CGD (collagen) (60), 2JE4 (HIV-1 protease dimer with inhibitor) (30), 1BRS (barnase) (57), 3G03 (MDM2) (61), 2NWD (lysozyme) (62), 4Q9G (FGF1) (63), and 2KID (sortase A) (64) were used. (B) Analytical data for the purified sequences of proinsulin, barstar, collagen, HIV-1 protease, MDM2^[1-118], lysozyme, FGF1, and sortase A*. For all cases, analytical HPLC data of the purified protein chains are presented as the main chromatographic trace with absorbance detection at 214 nm. The gradient for analytical HPLC was 5 to 65% B. A linear gradient of acetonitrile with 0.08% trifluoroacetic acid (TFA) added (solvent B) in water with 0.1% TFA added (solvent A) was used in all cases. Electrospray ionization (ESI) mass spectrum (upper left) and deconvoluted mass spectrum (upper right) are also shown in each case. Both spectra were obtained by summation of the entire LC peak; additional details on purification and analytical methods are in the SM section 5.

proteins accessed by AFPS therefore confirmed both the high substrate specificity and comparable activity to recombinant enzymes and literature values.

Binding studies of synthetic MDM2 and barnase confirmed specific affinities for their

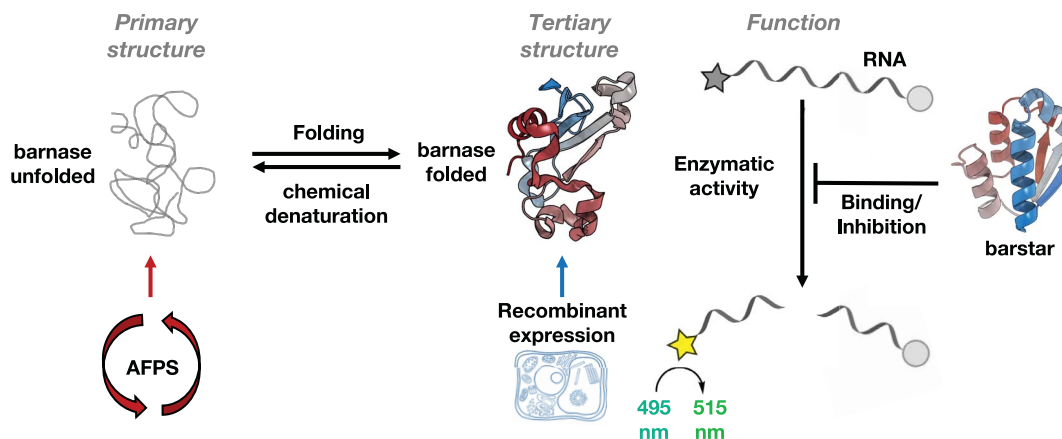
respective substrates. Barnase binds selectively and with high affinity to its inhibitor barstar. In a gel-based assay, recombinant barstar inhibited RNase activity of synthetic and recombinant barnase in a concentration-dependent manner (Fig. 4E) (47). In addition,

synthetic barstar obtained with AFPS performed comparably to recombinant barstar. To quantify binding of a synthetic protein to a known ligand, we also characterized the N-terminal binding domain of MDM2^[1-118] (32). The binding of MDM2 to p53 is a key

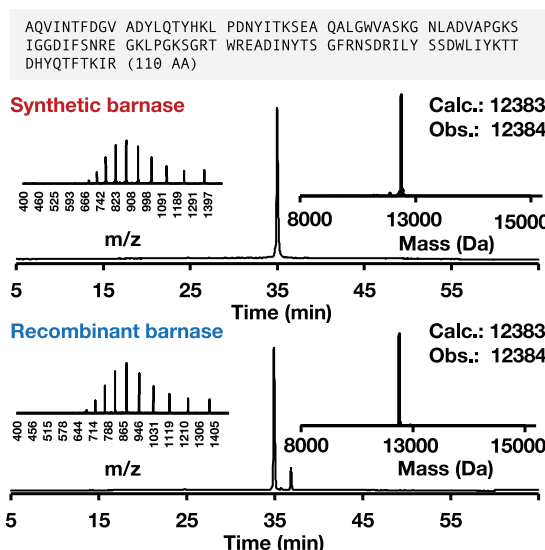
Fig. 4. Synthetic barnase and synthetic barstar fold into the native tertiary structure and display enzymatic activity comparable to recombinant samples.

(A) Conceptual overview of production and analysis methods. (B) Comparison of primary structures obtained from AFPS and recombinant expression. For both cases, analytical HPLC data of the purified barnase are presented as the main chromatographic trace with absorbance detection at 214 nm (additional details in the SM). ESI mass spectrum and deconvoluted mass spectrum of the purified peptide samples are displayed in the upper-left and the upper-right insets, respectively. Both spectra were obtained by summation over the entire LC peak in the chromatogram. (C) Structural evaluation of barnase in a chemical denaturation assay using urea as denaturant performed in triplicate; results are reported as mean $\pm$ SE. Error bars on the graph indicate SE. (D) Quantitative enzymatic activity assay performed in triplicate; error bars are not displayed for clarity. Details are outlined in the SM. k_{cat}/K_M values are reported as mean $\pm$ SE. (E) Barnase inhibition and binding assay using recombinant and synthetic barstar. 3.4 nM barnase was used in all conditions. Details are outlined in the SM. PDB 1BRS (barnase) (57) and 1AY7 (barstar) (59) were used.

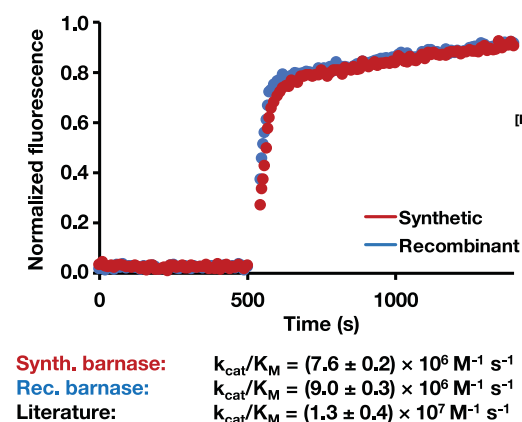
A Production, structure and function of barnase



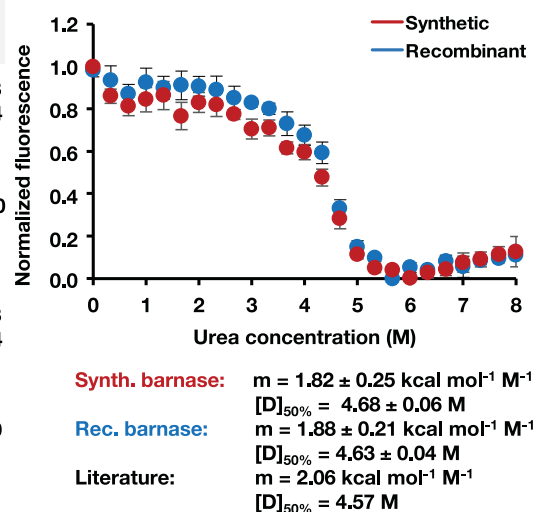
B Analysis of primary structure



D Fluorogenic activity assay



C Chemical denaturation of synthetic barnase



E Barstar inhibits barnase activity

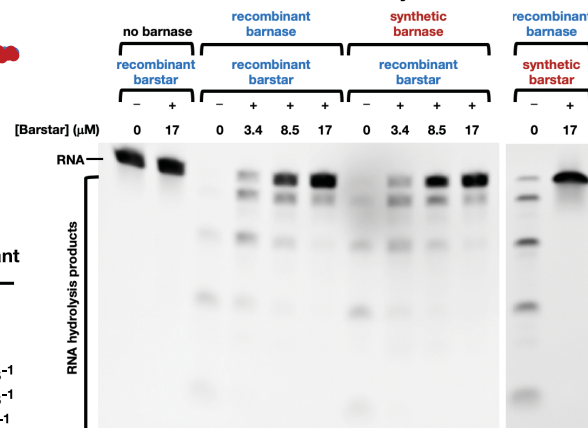
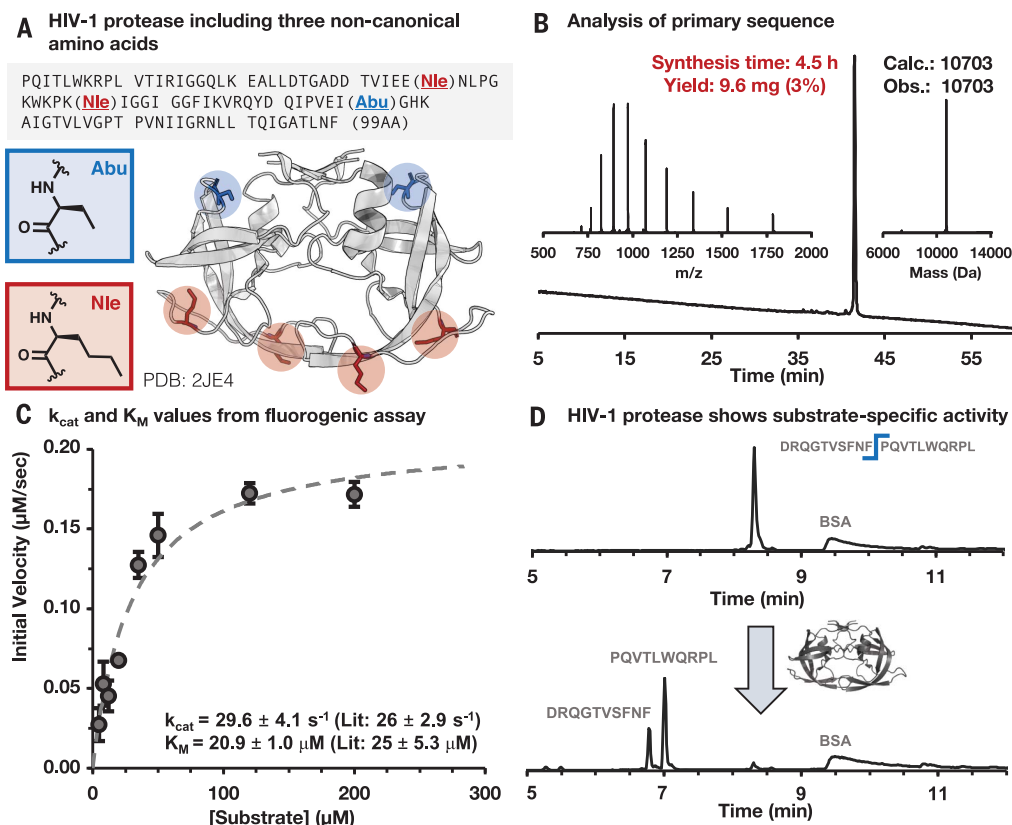


Fig. 5. Synthetic HIV-1 protease containing three noncanonical amino acids folds into the native dimer structure and displays enzymatic activity and substrate specificity comparable to literature samples. (A) Crystal structure of HIV-1 protease dimer with highlighted noncanonical amino acids aminobutyric acid (Abu, blue) and norleucine (Nle, red). PDB 2JE4 (HIV-1 protease dimer with inhibitor) (30) was used.

(B) Primary structure obtained from AFPS. Analytical HPLC data of the purified HIV-1 protease is presented as the main chromatographic trace with absorbance detection at 214 nm (additional details in the SM). ESI mass spectrum and deconvoluted mass spectrum of the purified sample are displayed in the upper-left and the upper-right insets, respectively. Both spectra were obtained by summation over the entire LC peak in the chromatogram. (C) Quantitative enzymatic activity assay performed in triplicate for the determination of k_{cat} and K_M values. Results are reported as mean $\pm$ SE. Error bars on the graph indicate SE. Lit., literature.

(D) Qualitative substrate specificity assay with model substrate p12nt, in which HIV-1 protease exclusively cleaves at a single Phe-Pro site whereas bovine serum albumin (BSA) stays intact; details are outlined in the SM.



interaction in multiple pathways up-regulated in cancer (48, 49). We folded milligram quantities of synthetic MDM2^[1-118] and characterized its binding to immobilized p53^[14-29] using biolayer interferometry (figs. S24 and S25). Synthetic MDM2^[1-118] displayed an affinity toward p53 [dissociation constant (K_d) = 6.25 μ M] comparable to the literature value (K_d = 5.45 μ M) obtained under the same folding conditions.

Discussion

The optimized AFPS protocol demonstrates advantages of flow chemistry over common batch methods, yielding peptide chains more than three times longer than previously accessible by routine standard SPPS (6). An improvement to existing flow protocols was achieved by rapid screening of variables in a reproducible reaction setup. Even though in this study AFPS yields superior results over traditional SPPS methods in terms of total synthesis time and crude product quality, general challenges associated with peptide synthesis, such as low atom economy and the use of DMF as a solvent, remain unsolved. A potentially limiting feature of our setup is synthesis scale. The capacity of the reactor used in our study allows up to 200 mg of resin with a loading of 0.49 mmol/g. Increased production output

can be achieved by incorporating a larger reactor in the current system, but such a modification will likely require specific optimization, toward which we performed preliminary investigations (19). Since we implemented AFPS, we have produced more than 5000 peptides and automatically collected in-line analysis data for all syntheses. Moving forward, this extensive, high-quality dataset could be leveraged to further improve peptide synthesis in flow using machine learning and other computational methods. Ultimately, we intend for this report to serve as a blueprint for the automated flow synthesis of other biopolymers and artificial sequence-defined polymers (50).

A robust, widely available routine method for chemical production of proteins is poised to have a strong impact on chemical biology and the development of new therapeutics. Our advances provide a viable solution to reliably assemble long linear peptide chains, shifting the focus in the field of chemical protein synthesis to improving folding protocols and, most importantly, applications. Combined with chemical ligation, rapid stepwise production of single-domain proteins by AFPS technology will extend the practical applications of total chemical synthesis to the majority of human proteins (those with a mass of up to ~30 kDa) (10, 51). In this respect, we envisage adapting

to our AFPS protocol the incorporation of peptide hydrazides for thioester-based ligation, an approach previously achieved with manual flow instrumentation (52). Additional research avenues opened by our method include rapid access to mirror-image proteins, posttranslationally modified proteins, and de novo-designed, abiotic proteins. Introduction of noncanonical amino acids as point mutations in native proteins will make accessible variants with considerably altered biological function, for example, catalytic activity (53, 54). Finally, AFPS has the potential to enable on-demand production of time-sensitive and potentially life-saving personalized medicine, such as for enzyme replacement therapy or neoantigen cancer vaccines (55, 56).

REFERENCES AND NOTES

1. J. Britton, C. L. Raston, *Chem. Soc. Rev.* **46**, 1250–1271 (2017).
2. D. Webb, T. F. Jamison, *Chem. Sci.* **1**, 675–680 (2010).
3. S. Bondalapati, M. Jbara, A. Briki, *Nat. Chem.* **8**, 407–418 (2016).
4. J. W. Chin, *Nature* **550**, 53–60 (2017).
5. R. B. Merrifield, *J. Am. Chem. Soc.* **85**, 2149–2154 (1963).
6. S. B. H. Kent, *Chem. Soc. Rev.* **38**, 338–351 (2009).
7. A. A. Zompra, A. S. Galanis, O. Werbitzky, F. Albericio, *Future Med. Chem.* **1**, 361–377 (2009).
8. P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **266**, 776–779 (1994).
9. V. Agouridas et al., *Chem. Rev.* **119**, 7328–7443 (2019).
10. P. E. Dawson, S. B. H. Kent, *Annu. Rev. Biochem.* **69**, 923–960 (2000).
11. A. El-Faham, F. Albericio, *Chem. Rev.* **111**, 6557–6602 (2011).

12. N. Ahmed, *Chem. Biol. Drug Des.* **91**, 647–650 (2018).
13. C. P. Gordon, *Org. Biomol. Chem.* **16**, 180–196 (2018).
14. A. J. Mijalis, A. Steinauer, A. Scheppartz, B. L. Pentelute, in *Flow Chemistry in Organic Synthesis* (Georg Thieme Verlag, Stuttgart, 2018), pp. 381–398.
15. E. Bayer, G. Jun, I. Halász, I. Sebestian, *Tetrahedron Lett.* **11**, 4503–4505 (1970).
16. T. J. Lukas, M. B. Prystowsky, B. W. Erickson, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 2791–2795 (1981).
17. A. J. Mijalis *et al.*, *Nat. Chem. Biol.* **13**, 464–466 (2017).
18. S. K. Mong, A. A. Vinogradov, M. D. Simon, B. L. Pentelute, *ChemBioChem* **15**, 721–733 (2014).
19. M. D. Simon, “Fast flow biopolymer synthesis,” thesis, Massachusetts Institute of Technology, Cambridge, MA (2017).
20. S. Fuse, Y. Otake, H. Nakamura, *Chem. Asian J.* **13**, 3818–3832 (2018).
21. S. Fuse, Y. Mifune, T. Takahashi, *Angew. Chem. Int. Ed.* **53**, 851–855 (2014).
22. L. K. Spare, V. Laude, D. G. Harman, J. R. Aldrich-Wright, C. P. Gordon, *React. Chem. Eng.* **3**, 875–882 (2018).
23. E. T. Sletten, M. Nuño, D. Guthrie, P. H. Seeberger, *Chem. Commun.* **55**, 14598–14601 (2019).
24. M. Schnölzer, P. Alewood, A. Jones, D. Alewood, S. B. H. Kent, *Int. J. Pept. Protein Res.* **40**, 180–193 (1992).
25. T. Michels, R. Dölling, U. Haberkorn, W. Mier, *Org. Lett.* **14**, 5218–5221 (2012).
26. M. Mergler, F. Dick, B. Sax, P. Weiler, T. Vorherr, *J. Pept. Sci.* **9**, 36–46 (2003).
27. S. A. Palasek, Z. J. Cox, J. M. Collins, *J. Pept. Sci.* **13**, 143–148 (2007).
28. S. Luisier, M. Avital-Shmilovici, M. A. Weiss, S. B. H. Kent, *Chem. Commun.* **46**, 8177–8179 (2010).
29. J. Schneider, S. B. H. Kent, *Cell* **54**, 363–368 (1988).
30. E. C. B. Johnson *et al.*, *J. Am. Chem. Soc.* **129**, 11480–11490 (2007).
31. A. Brik, C.-H. Wong, *Org. Biomol. Chem.* **1**, 5–14 (2003).
32. J. D. Oliner *et al.*, *Nature* **362**, 857–860 (1993).
33. D. Bresson *et al.*, *J. Clin. Invest.* **116**, 1371–1381 (2006).
34. E. Gasser, C. P. Moutos, M. Downes, R. M. Evans, *Nat. Rev. Endocrinol.* **13**, 599–609 (2017).
35. L. Serrano, J. T. Kellis Jr., P. Cann, A. Matouschek, A. R. Fersht, *J. Mol. Biol.* **224**, 783–804 (1992).
36. C. B. Anfinsen, *Science* **181**, 223–230 (1973).
37. S. B. Kent, P. F. Alewood, *Curr. Opin. Chem. Biol.* **22**, viii–xi (2014).
38. R. W. Hartley, *Trends Biochem. Sci.* **14**, 450–454 (1989).
39. J. T. Kellis Jr., K. Nyberg, A. R. Fersht, *Biochemistry* **28**, 4914–4922 (1989).
40. J. R. Knowles, *Science* **236**, 1252–1258 (1987).
41. U. Arnold *et al.*, *J. Am. Chem. Soc.* **125**, 7500–7501 (2003).
42. B. R. Kelemen *et al.*, *Nucleic Acids Res.* **27**, 3696–3701 (1999).
43. M. V. Toth, G. R. Marshall, *Int. J. Pept. Protein Res.* **36**, 544–550 (1990).
44. H. Ton-That, G. Liu, S. K. Mazmanian, K. F. Faull, O. Schneewind, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12424–12429 (1999).
45. J. J. Ling, R. L. Policarpo, A. E. Rabideau, X. Liao, B. L. Pentelute, *J. Am. Chem. Soc.* **134**, 10749–10752 (2012).
46. R. L. Policarpo *et al.*, *Angew. Chem. Int. Ed.* **53**, 9203–9208 (2014).
47. A. A. Vinogradov, E. D. Evans, B. L. Pentelute, *Chem. Sci.* **6**, 2997–3002 (2015).
48. F. Touti, Z. P. Gates, A. Bandyopadhyay, G. Lautrette, B. L. Pentelute, *Nat. Chem. Biol.* **15**, 410–418 (2019).
49. P. Chène, *Mol. Cancer Res.* **2**, 20–28 (2004).
50. J.-F. Lutz, M. Ouchi, D. R. Liu, M. Sawamoto, *Science* **341**, 1238149 (2013).
51. S. J. Wheelan, A. Marchler-Bauer, S. H. Bryant, *Bioinformatics* **16**, 613–618 (2000).
52. M. D. Simon *et al.*, *ChemBioChem* **15**, 713–720 (2014).
53. J. C. Li, T. Liu, Y. Wang, A. P. Mehta, P. G. Schultz, *J. Am. Chem. Soc.* **140**, 15997–16000 (2018).
54. H. Xiao *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6961–6966 (2015).
55. R. J. Desnick, E. H. Schuchman, *Nat. Rev. Genet.* **3**, 954–966 (2002).
56. N. L. Truex *et al.*, *Sci. Rep.* **10**, 723 (2020).
57. A. M. Buckle, G. Schreiber, A. R. Fersht, *Biochemistry* **33**, 8878–8889 (1994).
58. Y. Yang *et al.*, *J. Biol. Chem.* **285**, 7847–7851 (2010).
59. J. Sevcik, L. Urbanikova, Z. Dauter, K. S. Wilson, *Acta Crystallogr. D Biol. Crystallogr.* **54**, 954–963 (1998).
60. J. Bella, B. Brodsky, H. M. Berman, *Structure* **3**, 893–906 (1995).
61. A. Czarna *et al.*, *Cell Cycle* **8**, 1176–1184 (2009).
62. T. Durek, V. Y. Torbeev, S. B. H. Kent, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 4846–4851 (2007).
63. X. Xia, L. M. Longo, M. Blaber, *J. Pharm. Sci.* **104**, 566–576 (2015).
64. N. Suree *et al.*, *J. Biol. Chem.* **284**, 24465–24477 (2009).
65. A. J. Mijalis, amijalis/afps-integration: AFPS-analysis: Code used to extract and analyze data from AFPS synthesis files. Zenodo (2020).

ACKNOWLEDGMENTS

We thank T. F. Jamison, H. U. Stilz, L. F. Iversen, K. Little, and D. Lundsgaard for productive discussions and administrative support. We acknowledge the participants of the 26th American

Peptide Symposium and the 8th Chemical Protein Synthesis Meeting, especially P. E. Dawson, R. T. Raines, and S. B. H. Kent, for helpful discussions and for suggesting additional experiments. We also thank E. D. Evans, F. W. W. Hartrampf, and R. L. Holden for careful proofreading of the manuscript. We are grateful to N. L. Truex for providing recombinant sortase A*. Finally, we acknowledge C. M. T. Hartrampf for designing Fig. 1A. **Funding:** Financial support for this project was provided by Novo Nordisk. A.S., A.E.C., and C.K.S. gratefully acknowledge support from the National Science Foundation Graduate Research Fellowship under grant no. 1122374; A.E.C. is additionally supported by an MIT Dean of Science Fellowship. **Author contributions:** N.H., T.E.N., and B.L.P. conceptualized the research; N.H., M.P., A.J.M., and S.L. optimized synthesis conditions; A.J.M. provided the software for UV data analysis, and M.D.S. built the AFPS used in this report; M.P., A.J.C., and C.J. performed comparison of AFPS with traditional SPPS methods; N.H., A.S., M.P., A.J.C., A.E.C., S.H., S.A., C.K.S., and A.J.Q. synthesized, purified, and analyzed protein samples; N.H., Z.P.G., and B.L.P. conceptualized folding and biological evaluation of the synthetic proteins; N.H., A.S., M.P., Z.P.G., A.J.C., and X.Y. performed biological evaluation and expression of recombinant proteins; N.H., Z.P.G., A.L., and B.L.P. wrote the manuscript with input of all coauthors. **Competing interests:** B.L.P. is a cofounder of Amide Technologies and Resolute Bio. Both companies focus on the development of protein and peptide therapeutics. A.J.M. and M.D.S. hold equity in Amide Technologies. The following authors are inventors on patents and patent applications related to the technology described: A.J.M., M.D.S., and B.L.P. are co-inventors on U.S. patent application 20170081358A1 (23 March 2017) describing methods and systems for solid-phase peptide synthesis; M.D.S. and B.L.P. are co-inventors on U.S. patents 9,868,759 (16 January 2018), 9,695,214 (4 July 2017), and 9,169,287 (27 October 2015) describing solid-phase peptide synthesis processes and associated systems. **Data and materials availability:** Code for analysis of the UV-vis traces obtained from the AFPS instrument is available in a GitHub repository (65). All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6494/980/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S25
Tables S1 to S14
References (66–71)
MDAR Reproducibility Checklist

10 February 2020; accepted 30 March 2020
10.1126/science.abb2491

OCEAN CIRCULATION

Strengthening of the Kuroshio current by intensifying tropical cyclones

Yu Zhang^{1*}, Zhengguang Zhang¹, Dake Chen^{2,3}, Bo Qiu⁴, Wei Wang¹

A positive feedback mechanism between tropical cyclones (TCs) and climate warming can be seen by examining TC-induced energy and potential vorticity (PV) changes of oceanic geostrophic eddies. We found that substantial dissipation of eddies, with a strong bias toward dissipation of anticyclonic eddies, is directly linked to TC activity. East of Taiwan, where TCs show a remarkable intensifying trend in recent decades, the ocean exhibits a corresponding upward trend of positive PV anomalies. Carried westward by eddies, increasing numbers of positive PV anomalies impinge on the Kuroshio current, causing the mean current to accelerate downstream. This acts in opposition to decreasing basin-scale wind stress and has a potentially important warming impact on the extratropical ocean and climate.

In a warming climate, very intense tropical cyclones (TCs; categories 4 and 5 on the Saffir-Simpson scale) are becoming more intense and frequent, as proposed by theoretical and modeling studies (1–4) and supported by observations (5, 6). This behavior raises the question of what feedbacks can be induced by TC changes on climate. Much prior work has been focused on TCs' interaction with the quiescent upper ocean. For example, the decay of cold ocean wakes produced by storm-driven mixing causes net heating of the upper ocean and has been suggested to have important effects on global climate (7, 8). The ocean is anything but quiescent or uniform, though, and is full of dynamical structures of different scales. Among them, rotating structures with spatial scales of order O (10 to 100 km), temporal scales of order O (100 days), and vertical extents of 1000 m or more, usually referred to as mesoscale eddies, are the most ubiquitous and energetic; they play an essential role in transporting material and energy, and they modulate large-scale ocean circulation (9–11).

Because mesoscale eddies are ubiquitous, their encounters with TCs are not rare. By modulating subsurface density as well as thermal conditions, these synoptic-scale structures have been found to greatly affect the evolution of a storm's strength, posing a big challenge for operational prediction of TC intensity (12–15). Meanwhile, TCs can also have a large impact on eddies through more than one mechanism and in different or opposite

ways. For example, in central areas of TCs, strong and positive wind stress curl dominates, producing divergent surface currents, forcing cool water upward from below, and elevating isopycnals as much as 100 m, in the same sense as the circulation of cyclonic eddies and the associated upward displacements of isopycnals. Hence, oceanic cyclonic eddies can be enhanced or generated (16–18) while their mirror images, anticyclonic eddies, can be weakened. In outer regions of TCs, the wind stress curl is negative and very weak. Although the effects are to strengthen anticyclonic eddies yet weaken cyclonic eddies, opposite to those inside TCs' cores, the amplitudes are much reduced. In the meantime, at the smallest scale satisfying geostrophy, mesoscale eddies are always close to geostrophic balance. If this balance is disturbed through processes such as the shoaling of isopycnals by strong storms, the flow will adjust itself back to a state of geostrophic balance. This process, known as

geostrophic adjustment, was first considered by Rossby in 1938 (19) and has since been investigated in a variety of contexts both linear and nonlinear (20–23). Common to all those studies is that the flows in question hold less energy at their end states than they do initially, and energy is dispersed in the form of inertial gravity waves (19, 24). It is therefore anticipated that eddies under the influence of strong storms may be perturbed and subsequently undertake adjustment processes that attenuate them (25).

Evidence continues to mount that both mechanisms described above exist, working together to reduce the strength of anticyclonic eddies but acting against each other in changing cyclonic eddies. Which process dominates, and what overall influence the TCs exert on an underlying eddy field, remains unclear. To this end, a survey of eddies' lifetime evolution—in particular, immediately after their meeting with TCs—is necessary. The advent of satellite altimetry has enabled global and synoptic mapping of eddies with prominent surface signatures (26). The later combination of data from two altimeters has provided a time-series record of global maps of ocean eddies at high spatial and temporal resolutions [DT-2014 daily “two-sat merged” gridded product provided by archiving, validation, and interpretation of satellite oceanographic data (AVISO) (27)]. On the basis of that record, an eddy-tracking data archive was constructed, providing at daily frequency the center, amplitude, polarity, radius, and rotational velocity for those eddies identified and tracked by altimetry (28). In addition, Argo floats have provided depth profiles of temperature and salinity, from which eddies' vertical structures can be extracted (29). Taking advantage of both AVISO and Argo data, a law for the universal structure of

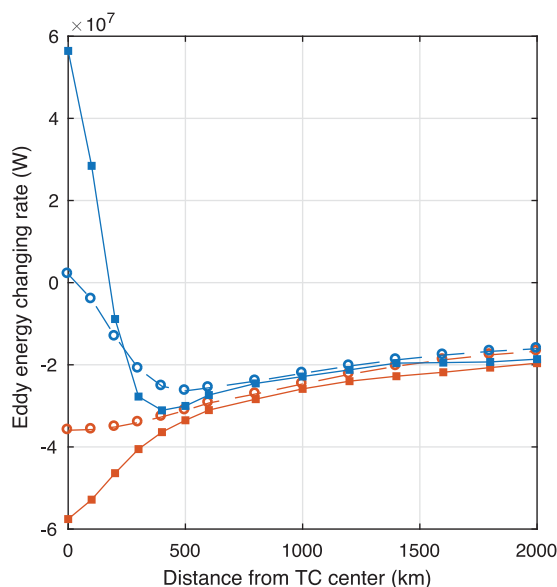


Fig. 1. Changing rate of eddy energy versus distance from TC center.

Mean changing rate of eddies' total energy is plotted against radial distance, averaged over 15 days after passing of TCs, for anticyclonic eddies (red curves) and cyclonic eddies (blue curves). Curves with circles are for all TCs in the western North Pacific Ocean between 1993 and 2014; those with squares are for more intense winds with speed greater than 40 m s⁻¹.

¹Physical Oceanography Laboratory/CIMST, Ocean University of China and Qingdao National Laboratory for Marine Science and Technology, Qingdao, P. R. China. ²State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou, P. R. China. ³Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, P. R. China.

⁴Department of Oceanography, University of Hawai'i at Manoa, Honolulu, HI, USA.

*Corresponding author. Email: yuz@ouc.edu.cn

eddies was uncovered (30), making it possible to accurately infer an eddy's three-dimensional structure from its surface signal. Using 20 years of these datasets in combination with TC observations in the western North Pacific, we carried out an analysis of the evolution of three-dimensional eddy energy and potential vorticity fields with a special focus on storm effects. [Potential vorticity (PV) is a dynamically conserved quantity depicting the tendency of a rotating fluid to spin.] Our results indicate that the interaction with TCs can introduce an important, newly recognized feature of the eddy field, which, together with the strengthening trend of TCs in a warming climate, has potentially strong influence on large-scale ocean circulation and climate.

Eddy response to TCs

Figure 1 shows the results of a survey of the rates of energy change of eddies during the 15-day period after their encounters with TCs in the western North Pacific Ocean (see supplementary materials). The rate of change depends on factors including storm wind speed, eddy polarity, and the radial distance, r , between an eddy and the storm's center at the time of meeting.

Cyclonic and anticyclonic eddies show obvious differences with little asymmetry about the storm track (see supplementary materials), especially under strong winds. Anticyclonic eddies that have passed within ~1000 km of the

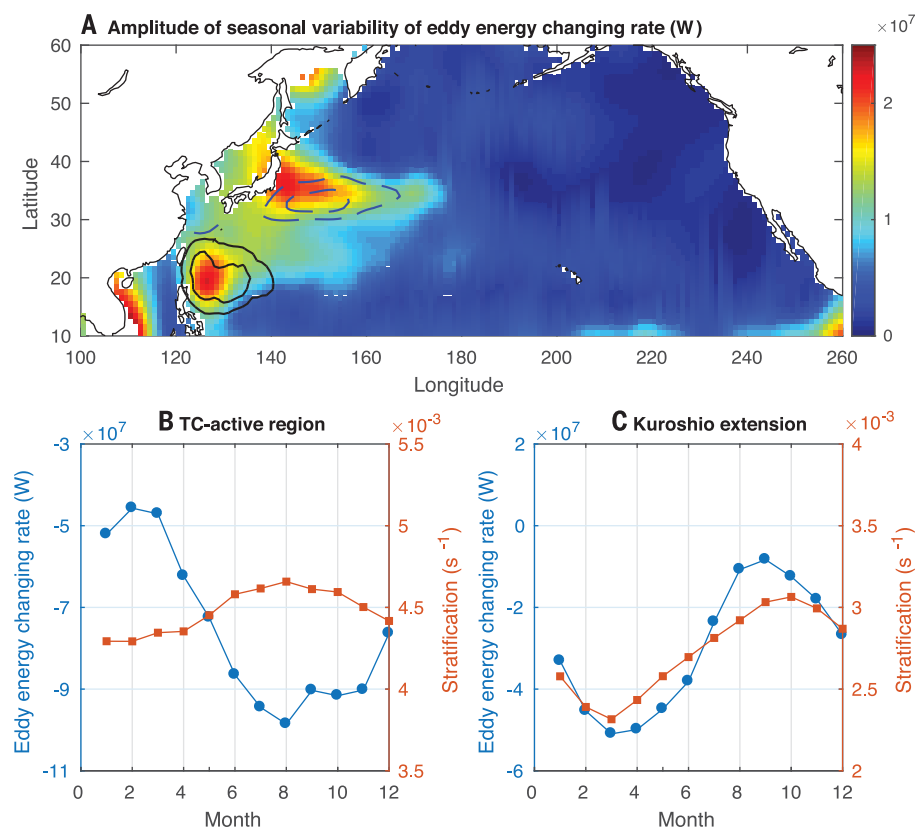
storm center generally experience substantial weakening afterward. On average, the energy decay rate declines from the storm center monotonically outward, asymptotically approaching the background dissipation level that is weaker than the central maximum value by a factor of 2 to 3 (red curves in Fig. 1). Cyclonic eddies' responses, although close to those of anticyclonic eddies' responses farther away from the storm, exhibit a tendency to strengthen progressively as they approach the storms' centers (blue curves in Fig. 1). Thus, TCs not only cause the entire eddy field to decay substantially more quickly than otherwise might have occurred but also make cyclonic eddies relatively stronger than anticyclonic eddies.

As seasonal phenomena, TCs develop mostly during the months of June through November, with peak activity occurring usually in late summer (31). If TCs' effect is dominant in driving eddy dissipation, the year-round energy evolution of eddies should exhibit similar seasonality.

In the North Pacific Ocean, only two areas demonstrate strong seasonal variation of the decay rates of eddies (Fig. 2A), and both are near the western boundary: One located at 20° to 30°N and 120° to 130°E, east of Taiwan, coincides with the region of the strongest TC activity (solid contours in Fig. 2A); the other one is located in the Kuroshio Extension. Both regions are prominent for vigorous eddy fields, but they differ from each other in two important aspects. First, their energetic eddy fields have

different origins. In the Kuroshio Extension region, a high eddy kinetic energy level arises from the meandering of the current itself (32), while along latitudes around 25°N, eddies are generated in the shearing flows between the eastward Subtropical Counter Current located in the 18° to 25°N band and the underlying westward North Equatorial Current (33). After generation, these eddies propagate mainly westward at speeds of a few centimeters per second, evolve as they move, enter and pass through the region of strong TC activity, and finally collide with the Kuroshio before they disappear (32). Second, seasonal variations of the eddies' decay rates in the two regions have different origins. In the TC-active region, the seasonal variation of the decay rate is in phase with that of TC activity (Fig. 2B), and so is the difference of decay rates between anticyclonic and cyclonic eddies (see supplementary materials). Both pieces of evidence indicate TCs' vital role in causing eddy dissipation. In the Kuroshio Extension region, the seasonal variation of the decay rate is almost 180° out of phase with that in the south, being strong in winter and spring but weak in summer and autumn (Fig. 2C). A closer look at different factors contributing to eddies' energy evolution reveals that the seasonality of the decay rate in this area is primarily due to the seasonally varying ocean stratification. Following the seasonal variation of solar irradiance, the buoyancy flux across the air-sea interface changes, as does

Fig. 2. Seasonal variation of eddy energy changing rate. (A) Seasonal variation amplitude of eddy energy changing rate in the North Pacific Ocean (color scale). TC intensity, defined as annual accumulated power input of TCs to the ocean, $\sum v^3$ (where v is the local instantaneous wind speed from the Best Track Dataset of Tropical Cyclones), is shown as solid contours. Seasonal variation amplitudes of oceanic stratification are shown as dashed contours. See supplementary materials for details. (B and C) Seasonal variation of eddy energy changing rate (blue) as compared with seasonal variation of oceanic stratification (red) in the TC-active region (120° to 130°E, 20° to 30°N) (B) and the Kuroshio Extension region (C). A 3-month running mean smoothing was applied to the time series.



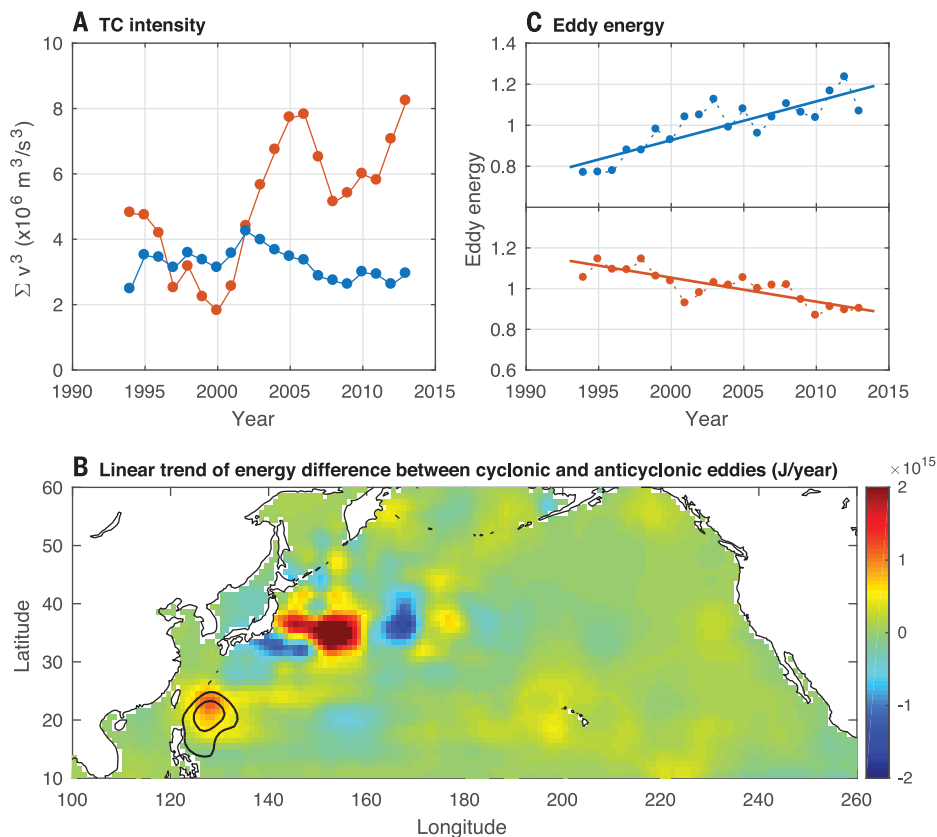


Fig. 3. Time series of TC intensity and energy difference between cyclonic eddies and anticyclonic eddies. (A) Time series of TC intensity, defined as annual accumulated power input in the TC-active area with winds below 40 m s^{-1} (blue) and above 40 m s^{-1} (red), from 1993 to 2014. (B) Linear trend of energy difference between cyclonic eddies and anticyclonic eddies in the North Pacific Ocean (color scale). Linear trend of TCs' annual accumulated power input is shown by contours. (C) Time series of normalized volume-integrated energy of cyclonic eddies (upper, blue dots) and anticyclonic eddies (lower, red dots) in the TC-active area. Linear trends are shown by solid lines. A 3-year running mean smoothing was applied to the time series. Each time series was normalized by dividing by its own 20-year mean.

the stratification in the ocean interior, which, as one of the most essential environmental factors shaping eddy characteristics, further modulates their energy and the associated evolution process. Near the tropics, both solar heating and ocean stratification change less across different seasons than they do at higher latitudes (dashed contours in Fig. 2A), whereas TC activity is much stronger, leaving TCs' seasonal forcing largely responsible for the seasonality of eddies' decay.

Our eddy energy analysis is performed in a Lagrangian framework (see supplementary materials): For each eddy identified by altimetry, the sum of kinetic and available potential energy is calculated on a daily basis along its track, and the 15-day low-pass-filtered changing rate is subsequently computed. More important, the volume integration instead of the surface component of energy is evaluated so as to include the effect of ocean stratification in driving the seasonality of the energy decay.

The TC-active region east of Taiwan (20° to 30°N , 120° to 130°E) not only has the strongest storms on average but also has displayed an obvious trend in TC intensity over the past decades, primarily due to intensified very strong storms (winds with speed above 40 m s^{-1} ; Fig. 3A). Because intense TCs are inclined to induce greater dissipation of anticyclonic eddies but to cause moderate to weak dissipation, or even discernible strengthening, of cyclonic

eddies, we anticipate relatively more energetic cyclonic eddies than anticyclonic eddies under intensifying TC activity. This expectation is confirmed by a trend analysis of energy difference between cyclonic and anticyclonic eddies in the entire basin. The result yields two regions of peak signals (Fig. 3B); the southern one located around 25°N displays a strong positive trend resulting from the strengthening of cyclonic eddies and the concurrent weakening of anticyclonic eddies (Fig. 3C). The region is nearly contiguous with the TC-active area (solid contours in Fig. 3B). Furthermore, its remarkable amplitude makes it distinct and isolated from both the neighboring areas and the latitudinal band farther to the east, from where eddies seen in this region are propagating. Therefore, the difference between the two types of eddies in this region is unlikely to be the result of any remote forcing. Instead, the strengthening of local TC activity introduces this new feature of the eddy field.

The Kuroshio Extension shows distinctly different trends between cyclonic eddies and anticyclonic eddies but with signs alternating zonally along the flow axis (Fig. 3B). Because eddies here are formed mostly by the pinch-off process of meanders, eddies generated from northward meanders are anticyclonic eddies with anomalous warm water at the core, and those generated from southward meanders

are cyclonic eddies with cold cores. In other words, north of the flow axis, anticyclonic eddies tend to dominate, whereas to the south cyclonic eddies prevail; however, with the meridional migration of the flow axis, the relative strength of the two types of eddies will be changed. For example, a region previously dominated by anticyclonic eddies will be filled with more cyclonic eddies as the axis moves to the north. Presumably the strong signals with alternating signs are related to spatial migration of the current axis, which is inhomogeneous in the zonal direction (34); these signals result from the redistribution of eddies but do not reflect net changes of the eddy field as a whole. Therefore, the only robust strengthening trend of cyclonic eddies relative to anticyclonic eddies in the North Pacific Ocean is the TC-driven trend east of Taiwan.

Strengthening of the Kuroshio current by eddies

The Kuroshio is primarily wind-driven. As the western boundary current of the North Pacific subtropical gyre, it acts as the return limb for the southward interior Sverdrup flow that is controlled by the basin-scale wind stress curl (32). Nonetheless, many other factors besides the basin-scale wind, such as the El Niño–Southern Oscillation (ENSO), the Pacific Decadal Oscillation (PDO), and the local wind stress and topography, influence it as well (35–37). In particular, eddy activity has

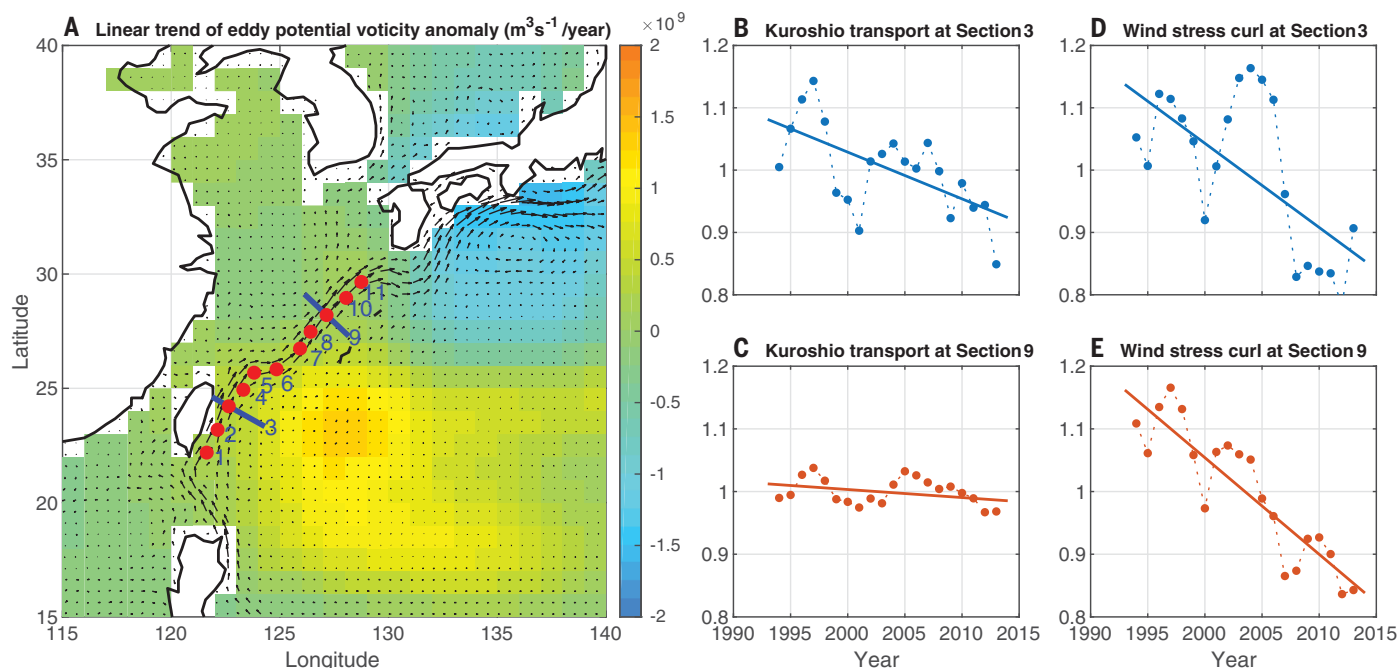


Fig. 4. Variation of Kuroshio transport along its path. (A) Linear trend of eddy potential vorticity (PV) anomaly in the western North Pacific Ocean (color scale). Mean velocity (arrow directions and lengths indicate direction and magnitude of velocity), coastlines (contour), and locations of cross sections (red dots) are shown. Sections 3 and 9 are denoted by solid blue lines. (B) Time series of normalized Kuroshio transport across section 3 (blue). (C) Same as (B) but for section 9 (red). (D) Time series of normalized strength of the wind

stress curl (blue) zonally averaged across the basin and meridionally within the latitudinal range of section 3 (20° to 25°N). (E) Same as (D) but for section 9 (25° to 30°N). Dots denote annual mean values with a 3-year running mean smoothing applied; solid lines denote linear trends. The Kuroshio transport across each section was calculated as the along-section integration of perpendicular surface geostrophic velocity. In (B) to (E), each time series was normalized by dividing by its own 20-year mean.

been suggested as a main modulator of the Kuroshio east of Taiwan (37, 38). Because this segment of the Kuroshio is also on the western edge of the TC-active region, a question naturally arises as to how the current is affected by the strengthening trend of cyclonic eddies relative to anticyclonic eddies discussed above.

The eddy flux of potential vorticity (PV) is the key to answering this question, because it comprehensively reflects the dynamical and thermodynamical links between the mean and eddy components of the flow field (39). A crucial difference between a cyclonic and an anticyclonic eddy is that the former has a positive PV anomaly while the latter has a negative one, both of which are carried westward like passive tracers by these fairly nonlinear and isolated eddies. Because cyclonic eddies are relatively stronger than anticyclonic eddies, the associated positive PV anomaly prevails (Fig. 4A), leading to a westward flux of net positive PV anomaly toward the mean current. According to the turbulent Sverdrup balance theorized by Rhines and Holland in 1979 (40), the westward positive eddy PV flux works to increase the PV of the mean component by accelerating the current northward (see supplementary materials). As shown below, this is the underlying mechanism for the

downstream acceleration of the Kuroshio in recent decades.

In the past two decades (1994–2013), Kuroshio transport east and downstream of Taiwan has undergone an obvious decreasing trend. The trend has been ascribed to the weakening of the subtropical wind stress curl (41), which, according to the classical Sverdrup balance, sets the transport of the western boundary current (42), but apparently this is not the whole story. For example, across both section 3 and section 9, the Kuroshio transport was noticeably weakened, but the upstream transport was reduced more remarkably: a 13% decrease at section 3 (Fig. 4B) versus a 3% decline at section 9 (Fig. 4C). Meanwhile, the strength of the basin-scale wind stress curl within this latitudinal band experienced a much greater decline than that of the Kuroshio transport, and the reduction amplitude slightly strengthened northward: a 27% decrease around the latitudes of section 3 (Fig. 4D) versus a more vigorous one, 31%, at section 9 (Fig. 4E). Therefore, some other factors must be accelerating the Kuroshio and counteracting the decelerating effects of the basin-scale wind.

To get a clearer picture, we take the difference between the two sections. The difference of Kuroshio transport reveals a large positive trend suggesting that the flow has

increased downstream in the past 20 years (Fig. 5A). The difference of the strength of the basin-scale wind stress curl, however, demonstrates a slight downward trend, corresponding to a tendency for downstream deceleration of the current (Fig. 5B). Hence, the classical Sverdrup balance depicting the dominant role of the basin-scale wind field cannot explain the observed acceleration of the current, and therefore, consistent with the turbulent Sverdrup balance theory, eddy PV flux is implicated.

Calculation of the eddy PV flux (see supplementary materials) between the two sections yields a curve (Fig. 5C) that is in good agreement with the downstream transport change (correlation coefficient of 0.69) in terms of both the interannual variability and the long-term upward trend. Moreover, at other locations along the current axis from south of Taiwan toward south of Japan (red dots in Fig. 4A), trends of the downstream change of Kuroshio transport and trends of the corresponding eddy PV flux (Fig. 5D) show a linear relation between the two that is consistent with our theoretical prediction.

All the above pieces of evidence point to the eddy effect as the reason for the increasing strength of Kuroshio transport despite the weakening basin-scale wind stress curl, indicating a linkage between TC activity and

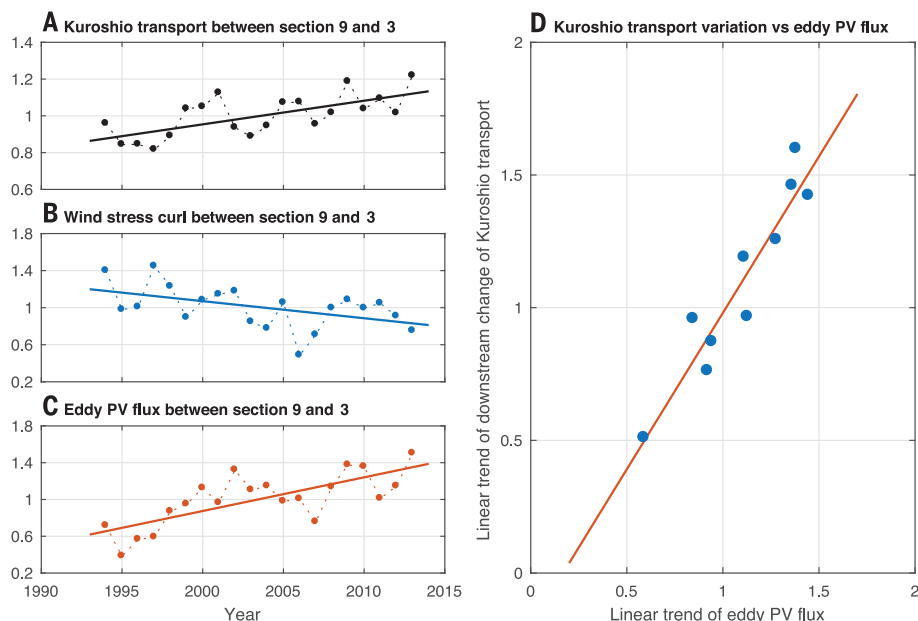


Fig. 5. Modulation of Kuroshio transport by eddy potential vorticity (PV) flux. (A) Time series of normalized Kuroshio transport difference between section 9 and section 3 (black). (B) Same as (A) but for the normalized strength of the wind stress curl zonally averaged across the basin (blue). (C) Same as (A) but for the normalized eddy PV flux (see supplementary materials) in the area bounded by the two sections (red). Dots denote annual

mean with a 3-year running mean smoothing applied; solid lines denote linear trends. (D) Scatterplot of linear trends of the normalized Kuroshio transport difference between section 2, 3, 4, 5, 6, 7, ..., 11 and section 1, and linear trends of the normalized eddy PV flux in corresponding areas. Linear fitting result is shown as a solid line. In (A) to (C), each time series was normalized by dividing by its own 20-year mean.

the Kuroshio transport change. Relative to TCs, oceanic eddies usually have much longer life spans and much slower translation speeds. Thus, eddies facilitate slow oceanic responses to TCs' more rapid forcing, which would further feed back onto the longer-term variations of the large-scale ocean circulation and climate change through eddy-mean flow interactions.

Conclusions

The kinetic energy of the ocean circulation is dominated by mesoscale eddies, which have been receiving much attention since they were first discovered in the early 1960s. Here, we have demonstrated a linkage between TC forcing and the substantial decay of the underlying eddy field, from which energy is transferred progressively downscale to smaller scales that are vulnerable to dissipation and mixing processes. Moreover, this decay process has pronounced seasonality due to the seasonal variation of TC occurrence, which paces the downscale energy cascade as well as the enhancement of the related mixing in the deep ocean. Finally, we have shown that the modification of the eddy field by TCs results in a positive feedback between TCs and climate change by modifying Kuroshio transport.

As the western boundary current of the North Pacific Ocean, the Kuroshio serves as a conduit for transporting warm water northward from its equatorial source region, thus

playing a crucial role in modulating the ocean and climate in mid- to high latitudes (43, 44). As a result of global warming in recent decades, the western Pacific Warm Pool has warmed notably, which may cause more heat to be carried northward by the Kuroshio, although to some extent this could be moderated by the weakening of Kuroshio transport. Global warming also may cause more intense TCs to occur at a higher frequency, which should increase the ratio of cyclonic eddies to anticyclonic eddies and thereby increase the trend of positive eddy-PV flux impinging onto the Kuroshio. That should accelerate downstream transport and contribute to further warming at higher latitudes. Notwithstanding the lack of a quantitative determination of the warming effect of TCs, it seems quite likely that overlooking this positive feedback mechanism could induce measurable bias in climate predictions. For a proper representation of eddies in climate models, more theoretical and modeling studies are needed to improve our understanding of the physical processes involved in the interactions among eddies, TCs, and large-scale ocean circulation.

REFERENCES AND NOTES

1. K. A. Emanuel, *Nature* **326**, 483–485 (1987).
2. G. J. Holland, *J. Atmos. Sci.* **54**, 2519–2541 (1997).
3. K. Oouchi et al., *J. Meteorol. Soc. Jpn.* **84**, 259–276 (2006).
4. K. A. Emanuel, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12219–12224 (2013).

5. K. Emanuel, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 12681–12684 (2017).
6. Y. K. Lim, S. D. Schubert, R. Kovach, A. M. Molod, S. Pawson, *Sci. Rep.* **8**, 16172 (2018).
7. K. Emanuel, *J. Geophys. Res.* **106**, 14771–14781 (2001).
8. M. F. Jansen, R. Ferrari, *Geophys. Res. Lett.* **36**, L06604 (2009).
9. R. Ferrari, C. Wunsch, *Tellus A* **62**, 92–108 (2010).
10. R. Ferrari, C. Wunsch, *Annu. Rev. Fluid Mech.* **41**, 253–282 (2009).
11. Z. Zhang, W. Wang, B. Qiu, *Science* **345**, 322–324 (2014).
12. B. Jaimes, L. K. Shay, *Mon. Weather Rev.* **136**, 4188–4207 (2009).
13. L. K. Shay, G. J. Goni, P. G. Black, *Mon. Weather Rev.* **128**, 1366–1383 (2000).
14. N. D. Walker et al., *Geophys. Res. Lett.* **41**, 7595–7601 (2014).
15. C.-C. Wu, C.-Y. Lee, I.-I. Lin, *J. Atmos. Sci.* **64**, 3562–3578 (2007).
16. L. Sun et al., *J. Geophys. Res. Oceans* **119**, 5585–5598 (2014).
17. X. Shang, H. Zhu, G. Gen, C. Xu, Q. Yang, *Adv. Meteorol.* **340432** (2015).
18. Z. Lu, G. Wang, X. Shang, *J. Phys. Oceanogr.* **46**, 2403–2410 (2016).
19. C.-G. Rossby, *J. Mar. Res.* **1**, 239–263 (1938).
20. A. E. Gill, *J. Fluid Mech.* **77**, 603–621 (1976).
21. H. W. Ou, *J. Phys. Oceanogr.* **14**, 994–1000 (1984).
22. A. C. Kuo, L. M. Polvani, *J. Fluid Mech.* **394**, 1–27 (1999).
23. A. Coutino, M. Stastna, *Nonlin. Processes Geophys.* **24**, 61–75 (2017).
24. A. Cahn Jr., *J. Meteorol.* **2**, 113–119 (1945).
25. B. Jaimes, L. K. Shay, *J. Phys. Oceanogr.* **40**, 1320–1337 (2010).
26. L.-L. Fu, *J. Geophys. Res.* **88**, 4331–4341 (1983).
27. N. Ducet, P. Y. Le Traon, G. Reverdin, *J. Geophys. Res.* **105**, 19477–19498 (2000).
28. D. B. Chelton, M. G. Schlax, R. M. Samelson, *Prog. Oceanogr.* **91**, 167–216 (2011).
29. J. Gould et al., *Eos* **85**, 179–184 (2004).
30. Z. Zhang, Y. Zhang, W. Wang, R. X. Huang, *Geophys. Res. Lett.* **40**, 3677–3681 (2013).

31. K. Emanuel, *Annu. Rev. Earth Planet. Sci.* **31**, 75–104 (2003).
32. B. Qiu, in *Encyclopedia of Ocean Sciences* (Academic Press, 2001), pp. 1413–1426.
33. B. Qiu, *J. Phys. Oceanogr.* **29**, 2471–2486 (1999).
34. B. Qiu, S. Chen, N. Schneider, *J. Clim.* **30**, 9591–9605 (2017).
35. C. Hwang, R. Kao, *Geophys. J. Int.* **151**, 835–847 (2002).
36. C.-R. Wu, *Prog. Oceanogr.* **110**, 49–58 (2013).
37. D. Zhang, T. N. Lee, W. E. Johns, C.-T. Liu, R. Zantopp, *J. Phys. Oceanogr.* **31**, 1054–1074 (2001).
38. Y.-C. Hsin, *J. Geophys. Res. Oceans* **120**, 1792–1808 (2015).
39. G. K. Vallis, *Atmospheric and Oceanic Fluid Dynamics* (Cambridge Univ. Press, 2006).
40. P. B. Rhines, W. R. Holland, *Dyn. Atmos. Oceans* **3**, 289–325 (1979).
41. Y.-L. Wang, C.-R. Wu, *Sci. Rep.* **8**, 14633 (2018).
42. J. Pedlosky, *Geophysical Fluid Dynamics* (Springer, ed. 2, 1987).
43. H. Bryden, S. Imawaki, in *Ocean Circulation and Climate*, G. Siedler, J. Church, J. Gould, Eds. (Academic Press, 2001), pp. 455–474.
44. Y.-O. Kwon *et al.*, *J. Clim.* **23**, 3249–3281 (2010).

ACKNOWLEDGMENTS

We acknowledge discussions with J. Pedlosky. **Funding:** Supported by National Natural Science Foundation grants 41722602, 41730535, and 41876007; National Basic Research Priorities Program of China grant 2015CB954300; and National Program on Global Change and Air-Sea Interaction grant GASI-IPOVAI-04.

Author contributions: Y.Z. and W.W. conceived the project and developed data analysis methodology; Y.Z., Z.Z., and W.W. carried

out data analyses; Y.Z. wrote the manuscript; and B.Q. and D.C. reviewed and edited the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The sources of the data used in this paper are listed in the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6494/988/suppl/DC1
Materials and Methods
Figs. S1 and S2
References (45–47)

7 April 2019; accepted 30 March 2020
10.1126/science.aax5758

SYNTHETIC BIOLOGY

Electrogenetic cellular insulin release for real-time glycemic control in type 1 diabetic mice

Krzysztof Krawczyk^{1*}, Shuai Xue^{1,2}, Peter Buchmann¹, Ghislaine Charpin-El-Hamri³, Pratik Saxena¹, Marie-Didiée Husherr¹, Jiawei Shao^{2,4}, Haifeng Ye², Mingqi Xie^{1,4}, Martin Fussenegger^{1,5†}

Sophisticated devices for remote-controlled medical interventions require an electrogenetic interface that uses digital electronic input to directly program cellular behavior. We present a cofactor-free bioelectronic interface that directly links wireless-powered electrical stimulation of human cells to either synthetic promoter-driven transgene expression or rapid secretion of constitutively expressed protein therapeutics from vesicular stores. Electrogenetic control was achieved by coupling ectopic expression of the L-type voltage-gated channel $Ca_v1.2$ and the inwardly rectifying potassium channel $K_{ir2.1}$ to the desired output through endogenous calcium signaling. Focusing on type 1 diabetes, we engineered electrosensitive human β cells (Electro β cells). Wireless electrical stimulation of Electro β cells inside a custom-built bioelectronic device provided real-time control of vesicular insulin release; insulin levels peaked within 10 minutes. When subcutaneously implanted, this electrotriggered vesicular release system restored normoglycemia in type 1 diabetic mice.

Precise control of dosage is essential for the success of any drug-based therapy (1–4). However, taking pills or administering biopharmaceuticals at regular intervals based on body weight, as is standard medical practice, is far from being precise and does not reflect the dynamics required for sophisticated metabolic interventions (1–4). Cell-based therapies capitalizing on implanted encapsulated designer cells engineered to fine-tune in situ production and systemic delivery of protein therapeutics in response to chemical and physical cues have shown promising results in proof-of-concept studies (5, 6). Because chemical control input is often limited, traceless physical cues such as light (optogenetics) (7–10) or heat [transmitted by magnetic fields (magnetogenetics) or radio waves (radiogenetics) (11–14)] are attractive for achieving rapid remote control of therapeutic transgene expression because they avoid the side effects of chemical trigger compounds (15, 16) as well as the challenges they may present with respect to bioavailability or pharmacodynamics (17–20). However, available physically triggered gene switches may require a high energy input (6, 7, 9), often involve complex chemical or inorganic cofactors (12, 21), and may require fine-tuning of the transcription of the ther-

apeutic transgenes, which slows down the overall response dynamics (5, 6, 9, 12, 22). Thus, direct cofactor-free wireless electrical stimulation of engineered cells to control vesicular secretion of protein therapeutics in a robust, adjustable, and repeatable manner would offer substantial advantages for medical applications by enabling direct communication between electronic devices and designer cells.

Although cellular metabolism and human-made electronics share similar operating principles in terms of input sensing, information processing, and output production, the core information transfer and processing functions of living and electronic systems are different, which limits their interoperability. Humans use ion gradients across insulated membranes to simultaneously process slow analog chemical reactions and communicate information in multicellular systems through soluble or gaseous molecular signals. In contrast, electronic systems use multicore central processing units to control the flow of electrons through insulated metal wires with gigahertz frequency and communicate information across networks via wired or wireless connections. Thus, direct electrical stimulation of gene expression or vesicular secretion requires a bioelectronic interface that manages electrical conduction between electrodes and electrosensitive designer cells, as well as conversion of electronic information via depolarization to protein production and release.

The first attempts to create an electrogenetic interface were reported more than a decade ago (23, 24), but that interface was neither direct nor usable under physiological conditions. More recently, a SoxR-based redox system that can control gene expression in *Escherichia coli* was reported (25), but this was also indirect and was too toxic for in vivo

application. Thus, despite decades of expertise in converting trigger-inducible bacterial and fungal repressor-operator interactions into synthetic mammalian gene switches, simple translation of bacterial electrogenetics into a mammalian cellular context has been unsuccessful because of the cytotoxicity, limited bioavailability, and poor clinical compatibility of electrosensitive redox compounds (23).

With the advent of optogenetics, it became possible to use illumination to control target gene expression remotely, and thus to indirectly link electrical stimulation via a light source with cellular transcription control (6, 10). This enabled glycemic control of experimental type 2 diabetes by controlling an optogenetic biomedical implant with a smartphone to upload instructions for designer cells to produce and systemically deliver a therapeutic dose of an insulinogenic peptide (6). However, the optogenetic device requires a considerable amount of energy to operate the light source (6, 10). The power efficiency associated with direct electrical stimulation is a major reason why clinically licensed pacemakers can be battery-powered for a lifespan of at least 15 years (26). Other major challenges to the clinical application of optogenetic technology include illumination-based cytotoxicity (27), the use of bacterial components (6, 10, 18–20), and the need for sophisticated chemical or inorganic cofactors that have side effects (28–30), poor bioavailability, or short half-lives in vivo (31). Other traceless physical control technologies based on electro-induced heat transmission, such as magnetotaxis and radiogenetics, share the same challenges (12, 21, 32, 33).

Diabetes is a common, chronic condition, and so is an attractive target for individualized precision treatment. Regulation of blood glucose levels is a closed-loop homeostatic process. Glucose-stimulated insulin release by pancreatic β cells involves uptake and metabolism of glucose, adenosine triphosphate-mediated closure of potassium channels, depolarization of the plasma membrane, and opening of the voltage-gated calcium channels, which results in an intracellular Ca^{2+} surge and concurrent rapid release of insulin from intracellular storage vesicles (34). For intervention in this process, we aimed to design a bioelectronic interface consisting of an implantable platform that combines electronics and electrosensitive designer cells that can release insulin on demand. The implant would incorporate a cell chamber containing semipermeable membranes that permit nutrient supply and product delivery via fibrous connective tissue, while protecting the designer cells from cellular host responses (35, 36) and securely containing them for safety reasons (37). To address this need, we describe here a direct cofactor-free electrogenetic interface

¹Department of Biosystems Science and Engineering, ETH Zurich, CH-4058 Basel, Switzerland. ²Shanghai Key Laboratory of Regulatory Biology, Institute of Biomedical Sciences and School of Life Sciences, East China Normal University, Shanghai 200241, People's Republic of China. ³Département Génie Biologique, Institut Universitaire de Technologie Lyon 1, F-69622 Villeurbanne Cedex, France. ⁴Key Laboratory of Growth Regulation and Transformation Research of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, People's Republic of China. ⁵Faculty of Science, University of Basel, CH-4058 Basel, Switzerland.

*Present address: Novartis Pharma AG, CH-4002 Basel, Switzerland.

†Corresponding author. Email: fussenegger@bsse.ethz.ch

to trigger vesicular secretion of insulin by using electrical stimulation to modulate the membrane polarization of human β cells engineered for ectopic expression of calcium and potassium channels ($\text{Electro}\beta$ cells). Furthermore, to validate our approach, we incorporated these electrosensitive designer cells into a bioelectronics implant and evaluated its performance in a mouse model of type 1 diabetes.

Membrane depolarization-based transcriptional control in mammalian cells

L-type voltage-gated calcium channels consist of α_1 , α_2 , δ , and β subunits and are essential for the functioning of cardiomyocytes, neurons, and endocrine cells (38). These channels open upon membrane depolarization, and the resulting calcium influx regulates muscle contraction, vesicular secretion of hormones, and NFAT (nuclear factor of activated T cells)-driven induction of target genes (39).

To design a mammalian transcription-control circuit responsive to membrane depolarization, we cotransfected human embryonic kidney (HEK) 293T cells with one of the three L-type voltage-gated calcium channels— $\text{Ca}_v1.2$, $\text{Ca}_v1.3_{\Delta 42A}$, or $\text{Ca}_v1.3_{\Delta 42}$ —encoded by the common α_2/δ_1 ($\text{pCa}_v\alpha_2\delta_1$, $\text{P}_{\text{hCMV}}\text{-}\alpha_2/\delta_1\text{-pA}$) and β_3 ($\text{pCa}_v\beta_3$, $\text{P}_{\text{hCMV}}\text{-}\beta_3\text{-pA}$) subunits and the respective channel-forming subunits α_1C ($\text{pCa}_v1.2$, $\text{P}_{\text{hCMV}}\text{-}\alpha_1C\text{-pA}$), α_1D_{42A} ($\text{pCa}_v1.3_{\Delta 42A}$, $\text{P}_{\text{hCMV}}\text{-}\alpha_1D_{42A}\text{-pA}$), or $\alpha_1D_{\Delta 42}$ ($\text{pCa}_v1.3_{\Delta 42}$, $\text{P}_{\text{hCMV}}\text{-}\alpha_1D_{\Delta 42}\text{-pA}$), as well as the reporter plasmid pMX57 encoding the human placental secreted alkaline phosphatase (SEAP) driven by the P_{NFAT3} promoter (pMX57, $\text{P}_{\text{NFAT3}}\text{-SEAP-pA}$) (Fig. 1A). Depolarization of channel-transgenic HEK-293T cells with 40 mM KCl revealed that ectopic expression of $\text{Ca}_v1.2$ showed the highest depolarization-triggered SEAP induction (Fig. 1B).

Coexpression of the inwardly rectifying potassium channel $\text{K}_{ir}2.1$ ($\text{pK}_{ir}2.1$, $\text{P}_{\text{hCMV}}\text{-}\text{K}_{ir}2.1\text{-pA}$), which has been reported to decrease the resting membrane potential of mammalian cells (40), substantially decreased basal SEAP expression and improved the overall induction profile of the depolarization-triggered $\text{Ca}_v1.2$ -mediated transcription control device (Fig. 1C). Combinatorial analysis of the importance of $\text{Ca}_v1.2$'s individual components for overall depolarization-triggered transcription control revealed that the channel-forming α_1C subunit was essential, whereas α_2/δ_1 and β_3 were not, although their absence reduced the maximum SEAP expression (fig. S1). Therefore, we used cells expressing the full $\text{Ca}_v1.2$ with the α_1C , α_2/δ_1 , and β_3 components as well as $\text{Kir}2.1$, referred to as ElectroHEK , in all follow-up experiments. Note that ElectroHEK cells are not activated by physiological ion concentrations, not even at life-threatening levels of KCl [6.5 mM (27)] or at CaCl_2 levels representing a medical emergency [3.5 mM (28)] (fig. S2).

Design and characterization of a synthetic electrogenetic mammalian transcription-control device

To test whether transgene expression could be directly triggered by electrically stimulated membrane depolarization, we used voltage-controlled square unipolar pulses with alternate polarization to electrostimulate the ElectroHEK cells transfected with the P_{NFAT3} -driven SEAP

expression vector (pMX57, $\text{P}_{\text{NFAT3}}\text{-SEAP-pA}$) (41–43) (Fig. 2A). Indeed, electric pulse stimulation triggered pMX57-transgenic ElectroHEK cells to produce high levels of SEAP (Fig. 2, B to D). The electrostimulated transgene expression could be fine-tuned by voltage (maximum SEAP induction at 50 V) (Fig. 2B) and could also be adjusted by altering the pulse length (maximum SEAP induction at 2 ms) (Fig. 2C).

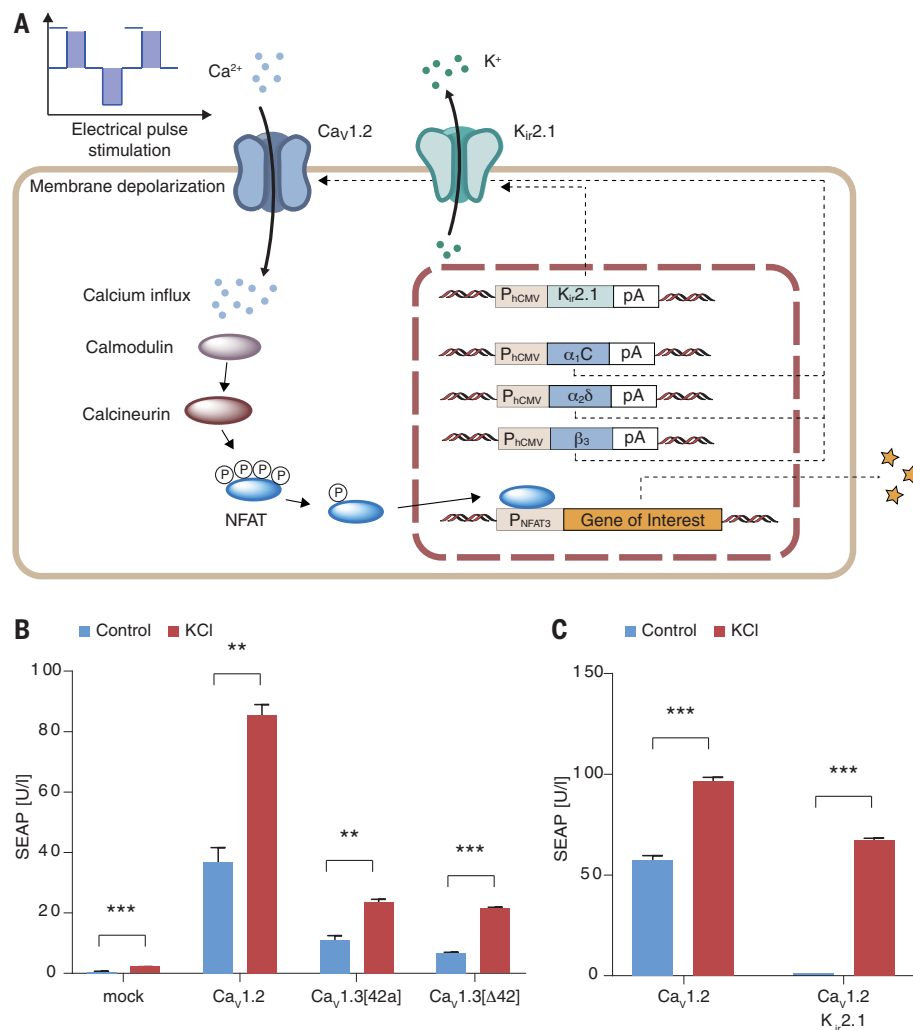


Fig. 1. Design of the electrogenetic circuit in mammalian cells. (A) Schematic representation of the electrogenetic circuit. The inwardly rectifying potassium channel lowers the resting membrane potential of HEK-293T cells, and electrical pulses depolarize the plasma membrane and open the L-type voltage-gated calcium channel. Calcium influx activates the calmodulin/calcineurin pathway, which leads to dephosphorylation of NFAT and its translocation to the nucleus, where it activates the NFAT-sensitive promoter and triggers transgene expression. (B) Comparative performance of three L-type voltage-gated calcium channels. Cells were cotransfected with P_{NFAT3} -driven SEAP reporter plasmid (pMX57), plasmids encoding α_2/δ_1 ($\text{pCa}_v\alpha_2\delta_1$, $\text{P}_{\text{hCMV}}\text{-}\alpha_2/\delta_1\text{-pA}$) and β_3 ($\text{pCa}_v\beta_3$, $\text{P}_{\text{hCMV}}\text{-}\beta_3\text{-pA}$), and one of the pore-forming subunits: α_1C ($\text{pCa}_v1.2$, $\text{P}_{\text{hCMV}}\text{-}\alpha_1C\text{-pA}$), α_1D_{42A} ($\text{pCa}_v1.3_{\Delta 42A}$, $\text{P}_{\text{hCMV}}\text{-}\alpha_1D_{42A}\text{-pA}$), and $\alpha_1D_{\Delta 42}$ ($\text{pCa}_v1.3_{\Delta 42}$, $\text{P}_{\text{hCMV}}\text{-}\alpha_1D_{\Delta 42}\text{-pA}$), to form $\text{Ca}_v1.2$, $\text{Ca}_v1.3_{\Delta 42A}$, and $\text{Ca}_v1.3_{\Delta 42}$, respectively. pcDNA3.1(+) was used as a mock plasmid. The cell membrane was depolarized with 40 mM potassium chloride (red bars); after 24 hours, SEAP was quantified in the supernatant. Blue bars show negative controls. (C) Coexpression of L-type voltage-gated calcium channel $\text{Ca}_v1.2$ and inwardly rectifying potassium channel $\text{K}_{ir}2.1$. Cells were cotransfected with $\text{pCa}_v1.2$ ($\text{P}_{\text{hCMV}}\text{-}\alpha_1C\text{-pA}$), $\text{pCa}_v\alpha_2\delta_1$ ($\text{P}_{\text{hCMV}}\text{-}\alpha_2/\delta_1\text{-pA}$), $\text{pCa}_v\beta_3$ ($\text{P}_{\text{hCMV}}\text{-}\beta_3\text{-pA}$), $\text{pK}_{ir}2.1$ ($\text{P}_{\text{hCMV}}\text{-}\text{K}_{ir}2.1\text{-pA}$), and pMX57 ($\text{P}_{\text{NFAT3}}\text{-SEAP-pA}$) in the molar proportions 1:1:1:1:3. Cells were depolarized with 40 mM KCl for 24 hours (red bars) and SEAP was quantified in supernatant samples. Data are means $\pm$ SEM; $n = 3$. ** $P < 0.01$, *** $P < 0.001$ (versus control).

Full activation of the system was reached after 4 hours of stimulation (Fig. 2D). Electrostimulation efficiency did not depend on the pulsing frequency within the range of 0.5 to 10 Hz (Fig. 2E). The parameter set used for effective electrostimulation did not decrease cell viability (fig. S3, A to D). Additionally, $\text{Ca}_v1.2$ -deficient HEK-293T cells were insensitive to electrostimulation (fig. S3E). Kinetic experiments revealed maximum SEAP expression 7 hours after the beginning of stimulation (fig. S4A) and confirmed the reversibility of the system (fig. S4B).

Design of the bioelectronic implant

Translation of electrostimulated gene expression into a clinical proof-of-concept bioelectronic implant required a more compact design for electrodes and electrostimulation. Simple miniaturization of the free-hanging electrodes used in the device described above did not provide efficient electrostimulation. Thus, we designed a custom-engineered cell culture insert containing electrodes on either side of a semipermeable membrane harboring a monolayer of electrosensitive ElectroHEK cells (Fig. 3A). Electrostimulation of pMX57 (P_{NFAT3} -SEAP-pA)-transfected ElectroHEK cells resulted in peak SEAP levels at 7.5 V (Fig. 3, B and C), which is

one order of magnitude lower than that of the previous free-hanging electrode arrangement, and at shorter pulse length (Fig. 3, D and E); both factors are important for high power efficiency of any electrostimulation device.

To enable electrostimulated transgene expression by electrosensitive cells in vivo, we designed a wireless-powered bioelectronic implant. The custom-engineered cell culture insert equipped with the electrodes was clicked into a 3D-printed FDA-licensed polyamide casing (Fig. 4, A and B) containing a sealed electronic switchboard (figs. S5 and S6) that generated the square unipolar pulses for electrostimulation of the encapsulated ElectroHEK cells. The implant's electronic circuitry was inductively powered and controlled by an extracorporeal field generator that wirelessly communicated with the bioelectronic implant at the ISM (industrial, scientific, and medical) frequency of 13.56 MHz (Fig. 4B and figs. S7 and S8). The voltage of the square pulses generated by the implant was dependent on the distance to the center of the field generator (fig. S9). The electronic circuit was insensitive to temperatures between 25° and 50°C (table S1). A control run of the bioelectronic implant validated wireless-controlled electrostimulated SEAP expression of pMX57-transfected ElectroHEK

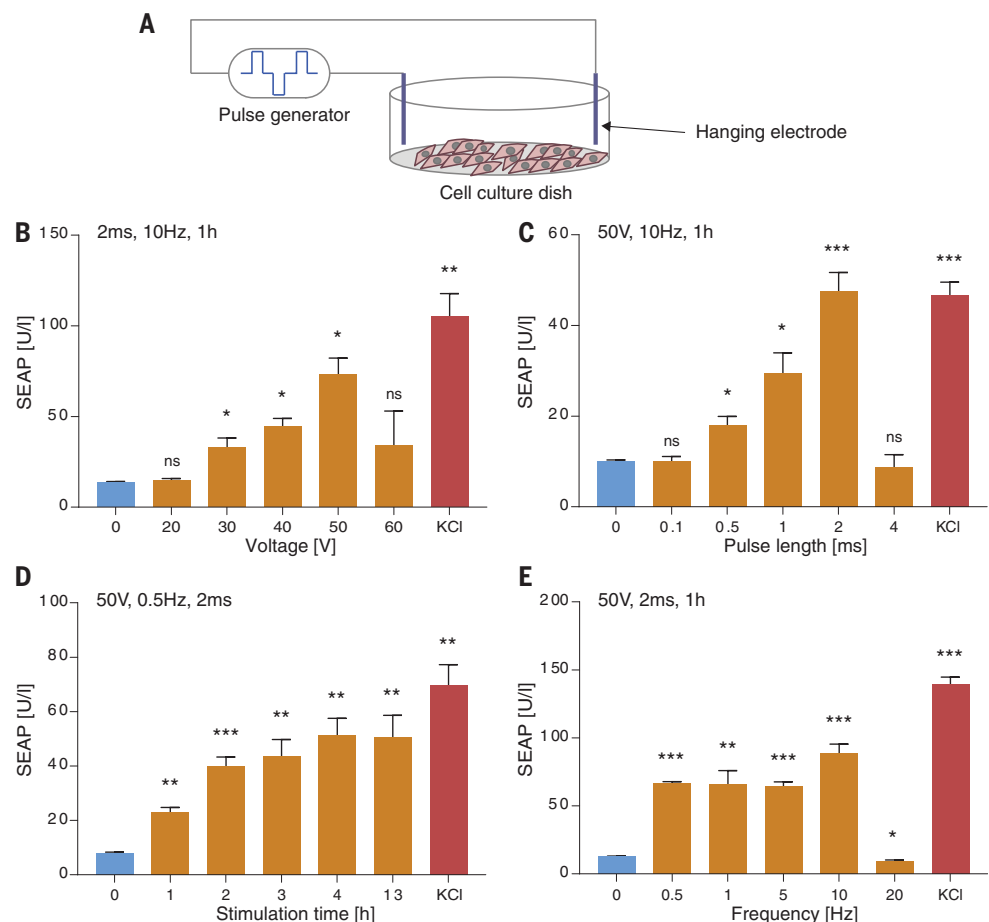
cells (Fig. 4C). We confirmed that the bioelectronic implants are rated IPX7 waterproof (International Protection Marking, IEC standard 60529) and show no cell leakage in a 5-day in vitro experiment (table S2).

$\text{Electro}\beta$ cells provide electrostimulated vesicular secretion

Because ElectroHEK -based insulin production is transcription-based, it lacks the rapid release dynamics of vesicular secretion characteristic of native pancreatic β cells (5). To engineer mammalian cells for electrostimulated vesicular release of insulin (Fig. 5A), we derived a monoclonal population, INS_{Vesc} , from the pancreatic β cell line 1.1E7 (44) by selection for deficiency in glucose sensitivity (Fig. 6, E and F), but with retention of the vesicular insulin secretion machinery. Indeed, electron micrographs of $\text{Electro}\beta$ cells, an INS_{Vesc} variant stably transgenic for constitutive expression of $\text{Ca}_v1.2$ and $\text{K}_{ir2.1}$ channels (pKK66 , $\text{P}_{\text{hEF1}\alpha}$ - $\alpha_1\text{C-P2A-K}_{ir2.1}$ -pA; pMX251, $\text{P}_{\text{hEF1}\alpha}$ - α_2/δ_1 -P2A- β_3 -pA) as well as Proinsulin-NanoLuc, a designer construct engineered to co-secrete insulin and the *Oplophorus gracilirostris* luciferase (NanoLuc) at an equimolar ratio in endocrine cell types (45) (Fig. 5A), revealed storage vesicles reminiscent of insulin-containing

Fig. 2. Characterization of the electro-genetic circuit in vitro. (A) Schematic representation of electrical stimulation setup. Cells were stimulated with carbon hanging electrodes producing monopolar pulses with alternate polarization.

(B to E) Cells were cotransfected with p $\text{Ca}_v1.2$ (P_{hCMV} - $\alpha_1\text{C}$ -pA), p $\text{Ca}_v\alpha_2\delta_1$ (P_{hCMV} - α_2/δ_1 -pA), p $\text{Ca}_v\beta_3$ (P_{hCMV} - β_3 -pA), pKK05 (P_{hCMV} - $\text{K}_{ir2.1}$ -pA), and pMX57 (P_{NFAT3} -SEAP-pA) in the molar proportions 1:1:1:1:3. SEAP assay was performed 24 hours after the beginning of the electrical stimulation procedure. Blue, orange, and red bars respectively denote unstimulated controls, electrically stimulated samples, and cells depolarized with 40 mM KCl. (B) Voltage dependence. Electrical stimulation was performed for 1 hour with 2-ms pulses at 10 Hz and the indicated voltage. (C) Pulse length effect. Electrical stimulation was performed for 1 hour at 10 Hz, 50 V, and the indicated pulse length. (D) Time course. Electrical stimulation was performed for the indicated period of time with 2-ms pulses at 0.5 Hz and 50 V. (E) Frequency effect. Electrical stimulation was performed for 1 hour with 2-ms pulses at 50 V and at the indicated frequency. Data are means $\pm$ SEM; $n = 3$. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ (versus control); ns, not significant.



granules of human islet-derived β cells (Fig. 5, D and E). Additionally, $\text{Electro}\beta$ cells showed well-correlated vesicular insulin and NanoLuc secretion in response to KCl-mediated (Fig. 5, B and C) or electrostimulated (Fig. 6, A and B) membrane depolarization. The stability and functionality of the $\text{Electro}\beta$ cell line were confirmed over at least 30 passages during 3 months in continuous culture (fig. S10).

We profiled the depolarization-based insulin release dynamics by electrostimulating $\text{Electro}\beta$ cells and recording the corresponding NanoLuc-mediated luminescence in the culture supernatant (Fig. 6C). Peak NanoLuc levels were reached within 10 min after electrostimulation (Fig. 6C), whereas transcription-based insulin production and secretion by ElectroHEK , $\text{HEK-}\beta$ (5), and OptoHEK cells (9) required 8 hours (fig. S11). When repeatedly electrostimulated, $\text{Electro}\beta$ cells recovered full secretory capacity after 4 hours (Fig. 6D). Most important, $\text{Electro}\beta$ cells did not show any glucose-sensitive insulin production, which ensures exclusive electrostimulation control of vesicular insulin secretion without interference from blood glucose levels (Fig. 6, E and F). Overall, $\text{Electro}\beta$ cells showed electrostimulation parameters similar to those of ElectroHEK cells (fig. S12, A to D). To illustrate the broad applicability of our approach, we also demonstrated electrostimulated vesicular secretion of glucagon by pancreatic alpha cells, which secrete the insulin counterregulatory hormone glucagon by calcium-triggered vesicular release (46) (fig. S13); this is akin to β cell-mediated insulin secretion.

Wireless electrostimulated vesicular secretion of insulin provides rapid glycemic control in type 1 diabetic mice

Native pancreatic β cells release the insulin stored in granules via a process known as vesicular secretion (34). The immediate release of stored insulin improves the response dynamics and rapidly restores blood glucose homeostasis in response to postprandial excursions. So far, designer cell-based proof-of-concept strategies to treat experimental diabetes have focused on transcriptional control, which is considered too slow to cope with postprandial blood glucose surges (5, 6, 12, 14, 21). For example, previously reported $\text{HEK-}\beta$ cells (5), which rely on transcriptional control and the classical secretory pathway for insulin release, require up to 24 hours to reach physiological blood insulin levels (fig. S14). Similar performance was observed for OptoHEK cells (9). In contrast, when placed into the wireless-powered bioelectronic implant (Fig. 4), $\text{Electro}\beta$ cells could reestablish postprandial glucose metabolism in insulin-deficient type 1 diabetic mice after a brief electrostimulation without causing hypoglycemic excursions (Fig. 7A) and could rapidly decrease blood

glucose levels to restore normoglycemia after electrostimulation (Fig. 7B). Notably, the results of glucose tolerance tests revealed comparable performance between $\text{Electro}\beta$ cells and human pancreatic islets, which are known to release insulin by vesicular secretion upon glucose sensing (fig. S7A). Fast vesicular secre-

tion was also confirmed by blood luminescence quantification (47), which showed a peak signal just 1 hour after electrostimulation, returning to baseline after 2 hours (Fig. 7C). Glycemia could also be controlled over longer periods of time without any sign of hypoglycemia (Fig. 7D).

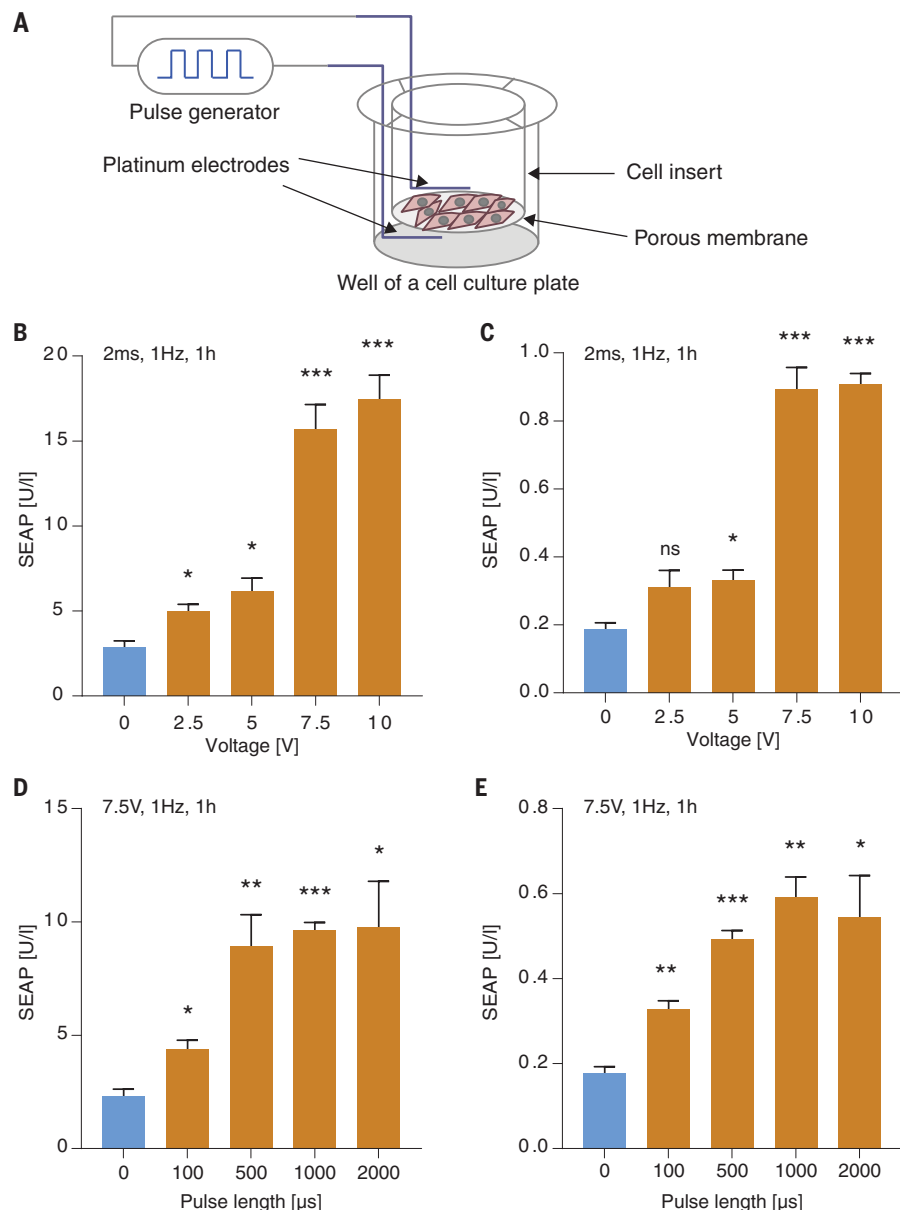


Fig. 3. Design and functionality of the bioelectronic implant in vitro. (A) Schematic representation of the stimulation setup in a cell culture insert. Two platinum electrodes (blue) were placed on opposite sides of the porous membrane covered with cells, and electrical pulse stimulation was applied. SEAP was quantified 24 hours after stimulation in the supernatants of the cell culture insert (above the membrane) and the well of the cell culture plate (below the membrane) to confirm that the secreted protein diffused across the membrane of the cell culture insert. (B and C) Voltage-dependent response of electrically stimulated pmX57-transfected $\text{Electro}\beta$ HEK cells grown in a cell culture insert. Cells were stimulated with 2-ms pulses at 1 Hz for 1 hour (orange bars). SEAP was measured in supernatant samples from the cell culture insert (above the membrane) (B) and from the cell culture well (below the membrane) (C). Blue bars denote negative controls. (D and E) Pulse length dependence. Cells were stimulated with 7.5 V pulses at 1 Hz for 1 hour (orange bars). SEAP was measured from supernatant samples from above (D) and below the cell layer (E). Blue bars denote negative controls. Data are means $\pm$ SEM; $n = 3$. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ (versus unstimulated control).

Biocompatibility and functional longevity of the bioelectronic implant

To validate the biocompatibility of the bioelectronic implants, we analyzed treated animals as well as explanted devices at 3 weeks after implantation, according to ISO 10993 (48), and we observed no material cytotoxicity, systemic kidney or liver toxicity, or alteration of hematologic profile or systemic immune responses; in addition, we saw no local immune-cell infiltration or substantial fibrotic tissue formation at the implant-tissue interface. There

was no apparent indication of implant-related cytotoxicity (fig. S15) or systemic toxicity (table S3), and no apparent difference in hematologic profiles among cell-containing and cell-free bioelectronic implants and biocompatible control implants (table S4). Likewise, we found no marked difference in the well-vascularized fibrous capsule surrounding the implants (fig. S16) or in immune-cell infiltration (fig. S17 and table S5) among cell-containing, cell-free, and biocompatible control implants. Mice implanted with $\text{Electro}\beta$ cell-containing

bioelectronic devices showed no change of body weight relative to untreated animals; also, signs of irritation or inflammation, as well as serum levels of inflammatory cytokines, were similar to or lower than those of animals treated with cell-free or biocompatible reference implants (figs. S18 and S19). Visual inspection of explanted bioelectronic devices showed no decomposition and no apparent erosion (fig. S20).

In view of the need for clinical translation toward a lifestyle-compatible therapeutic product, we adapted the bioelectronic implant architecture to allow repetitive exchange of individual cell batches over time (fig. S20A). Sequential in situ “refilling” of the implanted bioelectronic device with fresh batches of $\text{Electro}\beta$ cells without the need for surgical removal or replacement of the implant will reduce cost as well as implant-associated infections, while increasing patients’ convenience and treatment longevity. Insulin levels of type 1 diabetic mice, which had the $\text{Electro}\beta$ cells of their bioelectronic implants replaced once a week for a period of 3 weeks, were restored after remote-controlled electrostimulated insulin release by $\text{Electro}\beta$ cells (fig. S20, B and C). Together, these results suggest that the bioelectronic implant successfully integrates the advantages of electronics-based (49) and cell-based counterparts (5) and represents a promising approach to diabetes treatment.

Discussion

In this work, we have eliminated the need to use light as a converter between electronics and genetics, advancing optogenetics into electrogenetics by engineering a direct, cofactor-free electrogenetic interface that enables electronics to directly program gene expression as well as vesicular secretion in human cells. Furthermore, by incorporating electrogenetic designer cells ($\text{Electro}\beta$) containing this interface into a bioelectronic implant, we have successfully implemented a proof-of-concept device providing rapid electrostimulated insulin release for the treatment of experimental type 1 diabetes. The overall slow response dynamics associated with transcription-based control systems (5, 6, 9, 10, 12, 15, 16, 18–21) highlights the importance of vesicular secretion for the treatment of diabetes, which requires quick vesicular release of insulin to respond rapidly to postprandial blood glucose surges (50, 51). Indeed, we found that wireless electrical stimulation of vesicular insulin release from our engineered $\text{Electro}\beta$ cells encapsulated in a bioelectronic implant could attenuate postprandial hyperglycemia in type 1 diabetic mice with performance comparable to that of transplanted human pancreatic islets.

Taking account of the importance of economical manufacturing, we integrated all components of the bioelectronic implant into a 3D-printed polyamide casing. Although the

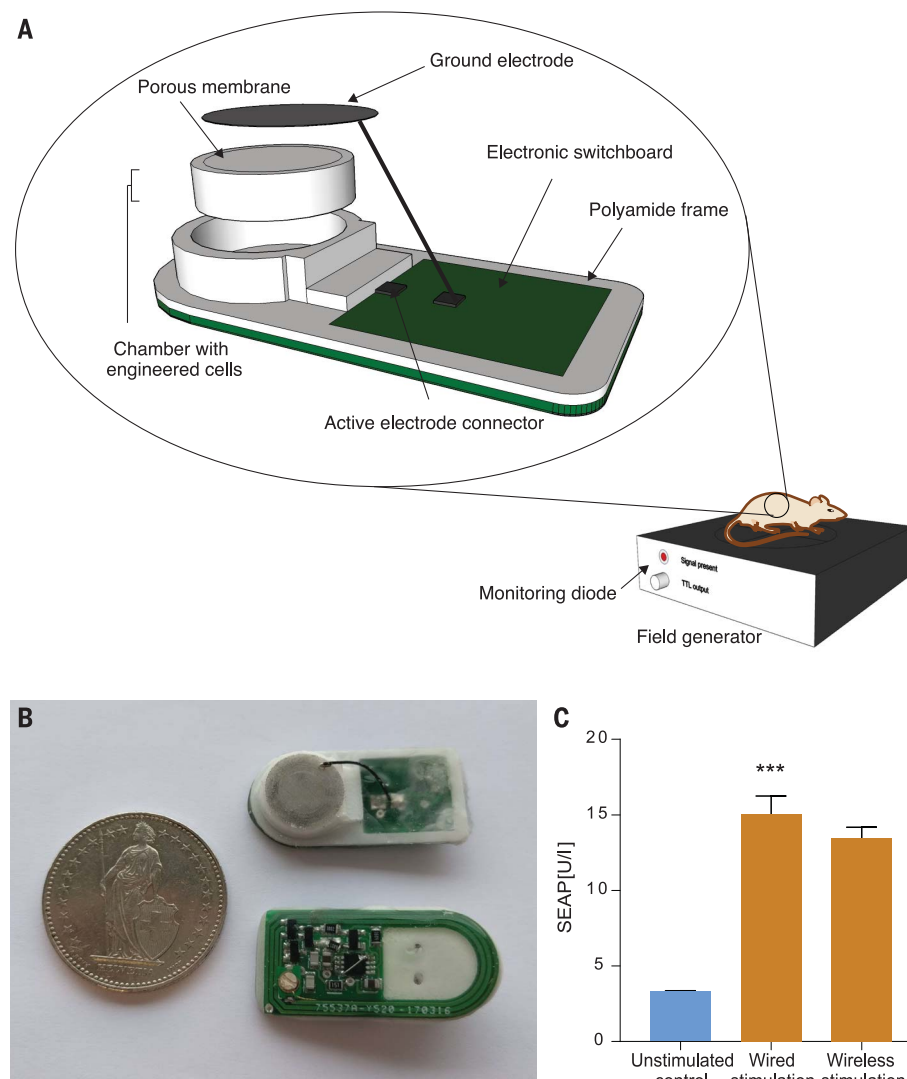
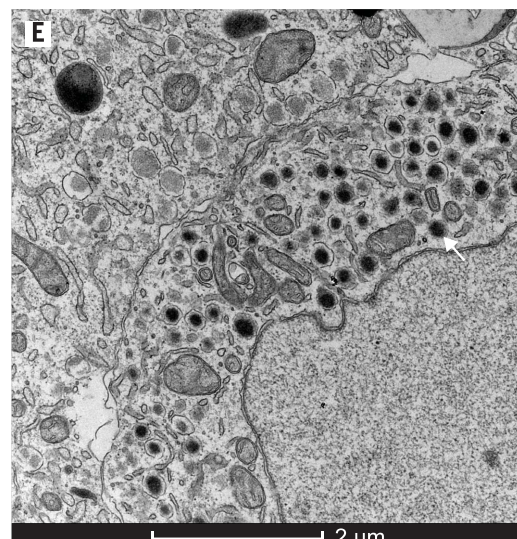
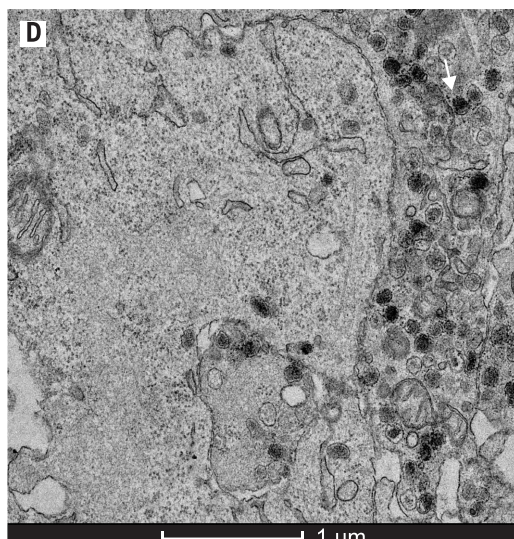
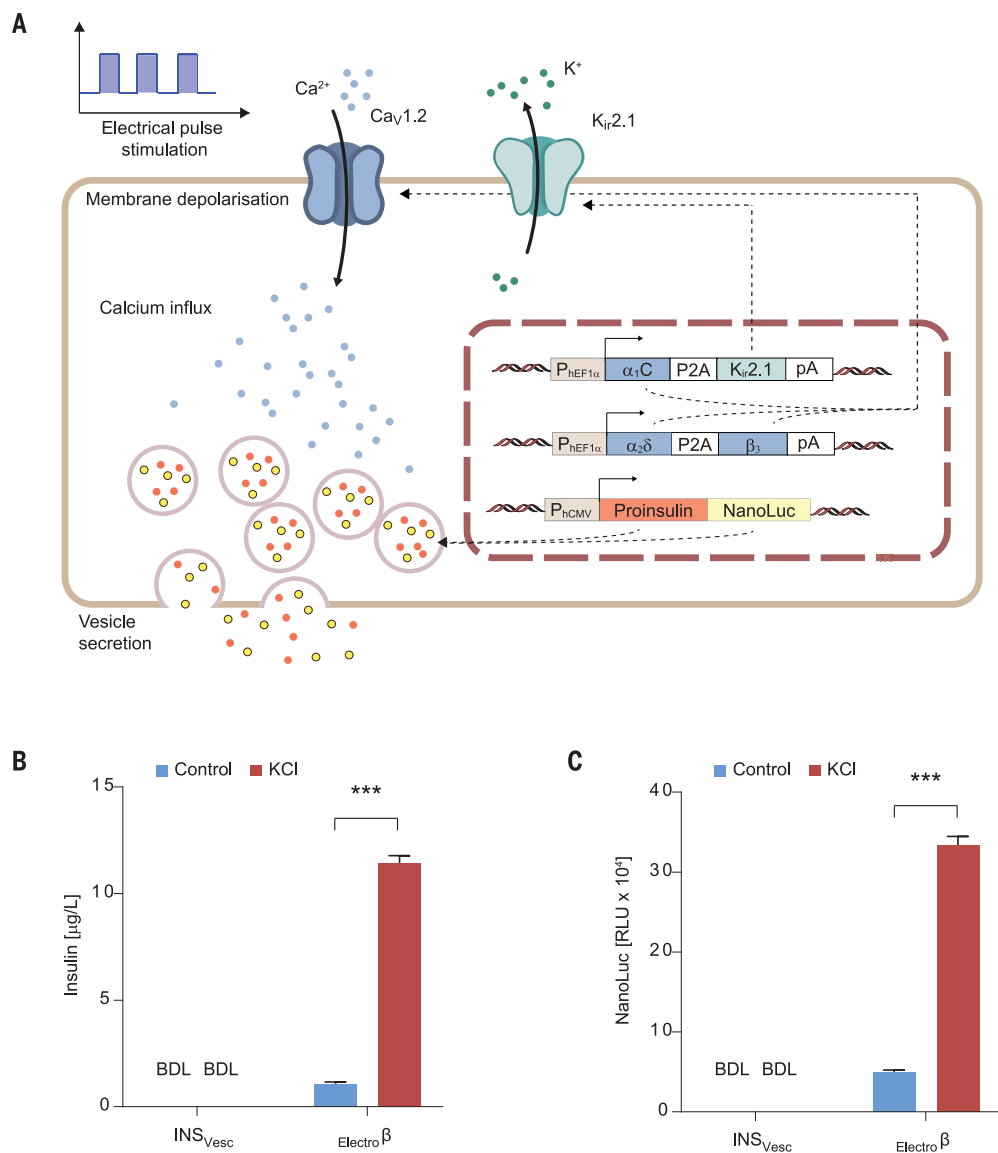


Fig. 4. Bioelectronic implant in vitro. (A) Three-dimensional model of a disassembled bioelectronic implant. A ring containing a porous membrane on one side can be assembled with a 3D-printed polyamide frame to form a cell chamber. The electronic switchboard is placed on the other side of the frame. The active platinum electrode (placed in the cell chamber; invisible in the model) is soldered to a connector on a switchboard. The ground electrode, made out of thin stainless steel mesh, is connected to the second connector on a switchboard. The bioelectronic implant can be placed subcutaneously on the dorsal side of the mouse, with the cell chamber facing down. The field generator provides wireless energy transmission. A red diode enables implant function monitoring. (B) Photograph of two bioelectronic implants with a coin (diameter 27.4 mm) for comparison. (C) Comparison of external generator-powered and implant-powered electrostimulation of pMX57-transfected $\text{Electro}\beta$ HEK cells. SEAP was measured in supernatant samples from above the cell layer. Data are means $\pm$ SEM; $n = 3$. *** $P < 0.001$ (versus control).

Fig. 5. Electrogenetic engineering of β cells. (A) Schematic representation of the electrically inducible insulin secretion pathway.

The inwardly rectifying potassium channel $K_{ir2.1}$ lowers the resting membrane potential, which keeps the voltage-gated calcium channel $Ca_v1.2$ closed. Electrical pulse stimulation causes membrane depolarization, opening of $Ca_v1.2$, and calcium influx, which stimulates vesicle secretion. Vesicles are loaded with pre-produced insulin (red dots) and NanoLuc (yellow dots).

(B and C) Comparison of insulin secretion by INS_{Vesc} and $Electro\beta$ cells. Vesicle secretion was quantified by insulin-specific enzyme-linked immunosorbent assay (ELISA) (B) and luminescence (C) before (blue bars) and after depolarization with 40 mM KCl (red bars). BDL, below detection limit. Data are means $\pm$ SEM; $n = 3$. *** $P < 0.001$ (versus control). (D) Transmission electron microscopy (TEM) image of $Electro\beta$ cells. White arrow indicates an insulin-containing vesicle. (E) TEM image of primary β cells from human pancreatic islets. White arrow indicates an insulin-containing vesicle.



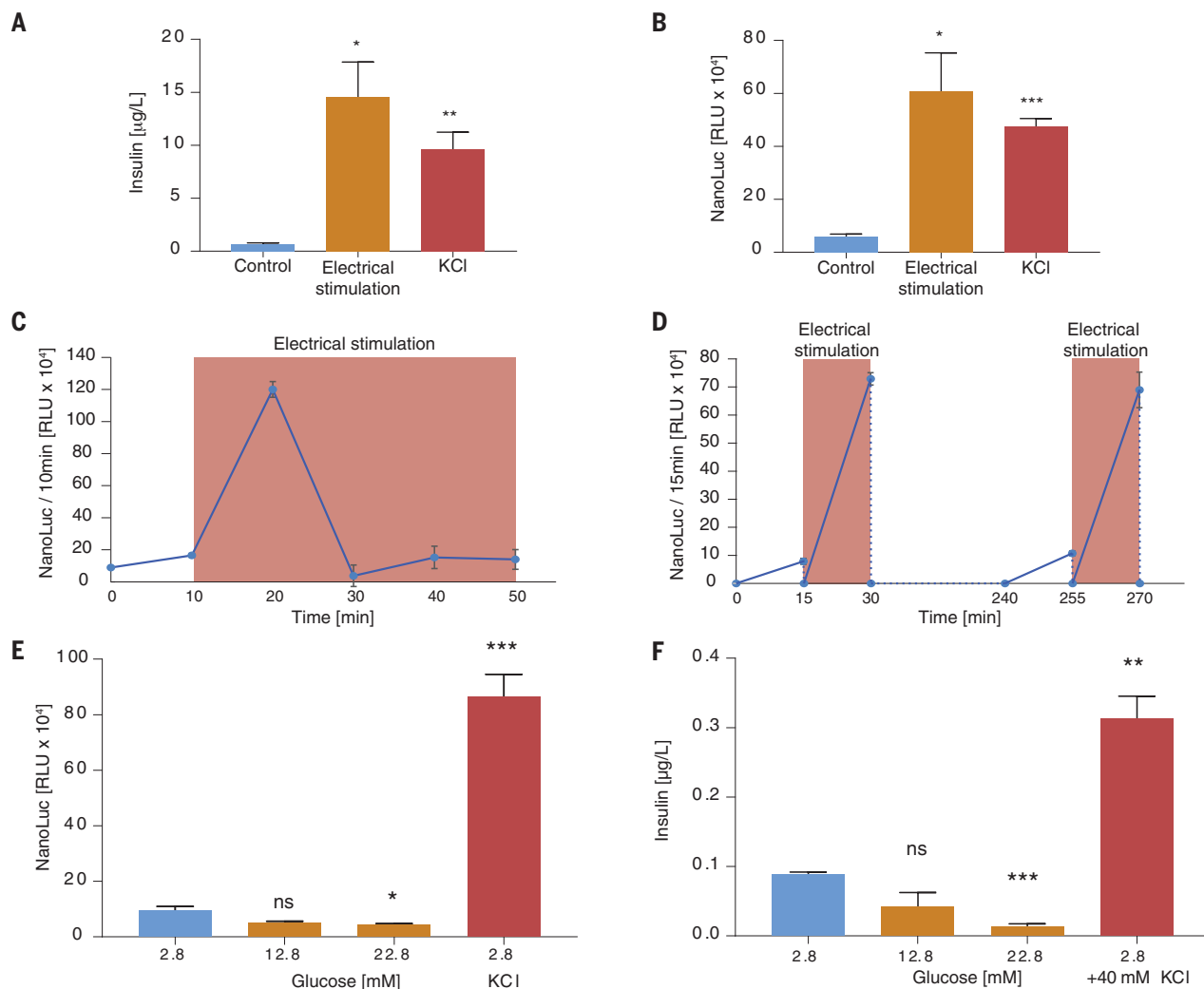


Fig. 6. Functionality of *Electroβ* cells in vitro. (A and B) Electrostimulation of *Electroβ* cells. Cells were seeded into cell culture inserts, and 24 hours later they were stimulated with electrical pulses (orange bars) or with 40 mM KCl (red bars). Blue bars denote negative controls. (A) Insulin content in the supernatant from inside the insert (above the cell layer) was measured by ELISA; $n = 3$. (B) Luminescence was measured in supernatant samples taken from inside the insert (above the cell layer); $n = 3$. (C) Secretion kinetics. *Electroβ* cells were seeded into cell culture inserts and stimulated with electrical pulses (red frame). Luminescence was measured in supernatant samples every 10 min; $n = 4$. (D) Reversibility assay. *Electroβ* cells were electrostimulated for 15 min twice, with

4-hour time intervals between the first and second electrostimulation; $n = 4$.

(E) Glucose-induced insulin release. *Electroβ* cells were incubated with various concentrations of glucose for 15 min (blue bar, 2.8 mM glucose; orange bars, elevated glucose; red bar, 2.8 mM glucose with 40 mM KCl). Luminescence was measured in supernatant samples; $n = 3$. (F) Glucose-induced insulin release. *INS_{Vesc}* cells were incubated with various concentrations of glucose for 60 min (blue bar, 2.8 mM glucose; orange bars, elevated glucose; red bar, 2.8 mM glucose with 40 mM KCl). Insulin content was quantified in supernatant samples; $n = 3$. Data are means $\pm$ SEM. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ (versus control).

bioelectronic implant could in principle be powered by batteries (52) (table S4), for practical reasons, including the limited space for implantation and the intrusiveness of animal experimentation, we chose to power the device inductively at 13.56 MHz, an FCC-licensed radio frequency that is reserved internationally for industrial, scientific, and medical devices and does not interfere with telecommunications. Because of the power efficiency of the implant, we speculate that wireless-powered control by wearable devices such as smartphones and smartwatches might be feasible in the near future.

However, reaching the full therapeutic potential of electrogenetics will require closed-loop control. Whereas classical medical interventions are open-loop, because the dose is largely determined by the physician on the basis of body weight, closed-loop systems enable feedback control that coordinates biomarker input to therapeutic output and provides an autonomous and self-sufficient interface with patients' metabolism. For electrogenetic type 1 diabetes control, this would mean using electronic blood glucose sensors to directly control electrostimulated insulin release in real time, much like the concepts currently being ex-

plored for prototypes of the bionic pancreas (53). However, electronic closed-loop systems operating in the bionic pancreas require frequent calibration and have a short lifespan of only a few days (49). On the other hand, incorporation of a microcontroller and/or a glucometer into our bioelectronic implant to achieve closed-loop insulin control should be a straightforward electrical engineering implementation. Most important, the delayed resorption of insulin from subcutaneous tissues to which insulin is delivered by the bionic pancreas requires dual-hormone control using glucagon to counteract or prevent

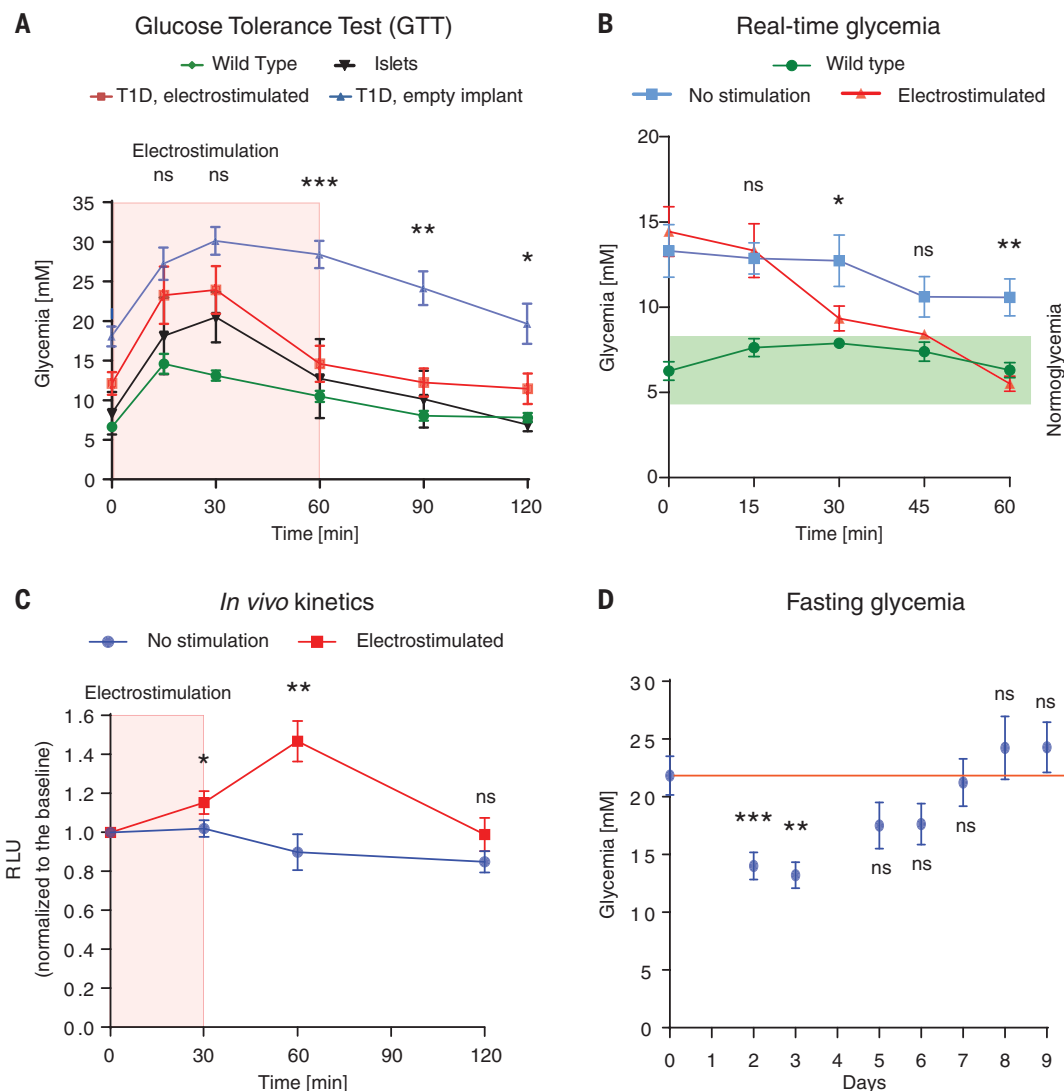


Fig. 7. Comparative analyses of $\text{Electro}\beta$ -containing bioelectronic implants in type 1 diabetic mice. Type 1 diabetic mice implanted on the back with $\text{Electro}\beta$ -containing bioelectronic devices were profiled for blood glucose dynamics. **(A)** Glucose tolerance test. At 48 hours after implantation, the $\text{Electro}\beta$ cells inside the bioelectronic implant were electrostimulated for 60 min (red line), then the animals were given intraperitoneal glucose injections and their blood glucose levels were monitored. All groups received intraperitoneal glucose injection (2 g per kg body weight). Wild type, $n = 8$; T1D, implant electrostimulated (type 1 diabetes, activated implant), $n = 6$; T1D, empty implant (type 1 diabetes, implant without cells), $n = 10$; islets (human pancreatic beta islets), $n = 3$. Statistical significance of differences between the electrostimulated and mock groups was calculated. **(B)** Real-time glycemia measurement. Fasted type 1 diabetic mice implanted with $\text{Electro}\beta$ -containing bioelectronic implants were electrostimulated for 30 min and their glycemic profile was recorded.

Nonstimulated control (T1D, implanted mice), $n = 6$; stimulated group (T1D, implanted mice), $n = 7$; wild-type controls, $n = 6$. The green frame indicates the normoglycemic range (4.4 to 7.2 mM). Statistical significance was calculated between electrostimulated mice and nonstimulated controls. **(C)** Blood luciferase kinetics of animals implanted with $\text{Electro}\beta$ cell-containing implants electrostimulated for 30 min (red line; $n = 6$). NanoLuc was quantified from microliter-scale blood samples every 30 min. The blue line indicates the nonelectrostimulated negative control ($n = 5$); red frame indicates electrostimulation time. Statistical significance of differences versus time point 0 was calculated with a paired t test. **(D)** Fasting glycemia. Type 1 diabetic mice were implanted with $\text{Electro}\beta$ -containing bioelectronic implants and fasting glycemia was recorded for more than 1 week. Orange line indicates the initial level of average glycemia. Statistical significance of differences versus time point 0 was calculated with a paired t test. Data are means $\pm$ SEM. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

insulin-mediated hypoglycemia (54, 55). We show here that glucagon can be released from pancreatic α cells by vesicular secretion, just as insulin is from β cells; this suggests that a dual-hormone electrogenetic system using two types of engineered cells would be feasible. Nonetheless, dual-hormone control is not expected to be necessary with our electro-

genetic system because, as noted above, the dynamics of electrostimulated vesicular insulin secretion from $\text{Electro}\beta$ cells appear to be comparable with those of human pancreatic islets. Furthermore, the demonstration that our system works in two different types of cells suggests broad potential applicability of electrogenetics for electrostim-

ulated hormone release in future cell-based therapies.

As in the case of the bionic pancreas (53), long-term functionality of cellular implants remains a major challenge in designing next-generation encapsulated cell-based therapeutic devices (56). A recent clinical trial using encapsulated pancreatic progenitor cells, the

precursor phenotype of insulin-secreting β cells (ViaCyte's VC-01), confirmed the need for further technological development to promote engraftment (57). Long-term functionality of cells inside implants remains among the challenges facing translation of academic proof-of-concept studies into clinical reality. In this context, the first initiatives to improve viability (Beta-O2 Technologies Ltd.; β Air) as well as vascularization of encapsulated cells (58) (e.g., ViaCyte's PEC Direct or Sernova's Cell Pouch System) have already begun in industry.

We have shown that wireless electrical stimulation of insulin release by electrosensitive designer cells inside a bioelectronic implant was able to rapidly restore normoglycemia in type 1 diabetic mice. The adoption of wireless electronic devices that can program the release of biopharmaceuticals, either via the secretory pathway or vesicular secretion, by means of direct communication between the device and implanted cells is expected to open up many new opportunities for advanced precision healthcare optimized for individuals.

REFERENCES AND NOTES

- M. M. Aye, S. L. Atkin, *Drug Healthc. Patient Saf.* **6**, 55–67 (2014).
- A. Goto, O. A. Arah, M. Goto, Y. Terauchi, M. Noda, *BMJ* **347**, f4533 (2013).
- G. P. Leese et al., *Diabetes Care* **26**, 1176–1180 (2003).
- C. Uduku, N. Oliver, *Curr. Opin. Pharmacol.* **36**, 29–33 (2017).
- M. Xie et al., *Science* **354**, 1296–1301 (2016).
- J. Shao et al., *Sci. Transl. Med.* **9**, eaal2298 (2017).
- T. Kushibiki, S. Okawa, T. Hirasawa, M. Ishihara, *Gene Ther.* **22**, 553–559 (2015).
- T. Kim, M. Folcher, M. Daoud-El Baba, M. Fussenegger, *Angew. Chem. Int. Ed.* **54**, 5933–5938 (2015).
- H. Ye, M. Daoud-El Baba, R. W. Peng, M. Fussenegger, *Science* **332**, 1565–1568 (2011).
- M. Folcher et al., *Nat. Commun.* **5**, 5392 (2014).
- H. A. Andersson, Y. S. Kim, B. E. O'Neill, Z. Z. Shi, R. E. Serda, *Vaccines* **2**, 216–227 (2014).
- V. Ortner et al., *J. Control. Release* **158**, 424–432 (2012).
- H. Huang, S. Delikanli, H. Zeng, D. M. Ferkey, A. Pralle, *Nat. Nanotechnol.* **5**, 602–606 (2010).
- S. A. Stanley, J. Sauer, R. S. Kane, J. S. Dordick, J. M. Friedman, *Nat. Med.* **21**, 92–98 (2015).
- M. Fussenegger et al., *Nat. Biotechnol.* **18**, 1203–1208 (2000).
- W. Weber et al., *Nat. Biotechnol.* **20**, 901–907 (2002).
- B. G. Auner, C. Valenta, J. Hadgraft, *J. Control. Release* **89**, 321–328 (2003).
- M. Gitzinger, C. Kemmer, M. D. El-Baba, W. Weber, M. Fussenegger, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 10638–10643 (2009).
- M. Gitzinger et al., *Nucleic Acids Res.* **40**, e37 (2012).
- H. Wang, H. Ye, M. Xie, M. Daoud El-Baba, M. Fussenegger, *Nucleic Acids Res.* **43**, e91 (2015).
- S. A. Stanley et al., *Science* **336**, 604–608 (2012).
- A. Prindle et al., *Nature* **508**, 387–391 (2014).
- W. Weber et al., *Nucleic Acids Res.* **37**, e33 (2009).
- W. Weber et al., *Nat. Biotechnol.* **22**, 1440–1444 (2004).
- T. Tschirhart et al., *Nat. Commun.* **8**, 14030 (2017).
- D. Katz, T. Akiyama, *Ann. Noninvasive Electrocardiol.* **12**, 223–226 (2007).
- J. H. Stockley et al., *Sci. Rep.* **7**, 849 (2017).
- T. D. Hinds Jr. et al., *J. Biol. Chem.* **291**, 25179–25191 (2016).
- Y. Uda et al., *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11962–11967 (2017).
- Z. Hu et al., *Int. J. Mol. Sci.* **16**, 22621–22635 (2015).
- H. M. Beyer et al., *ACS Synth. Biol.* **4**, 951–958 (2015).
- G. Jarockyte et al., *Int. J. Mol. Sci.* **17**, 1193 (2016).
- B. A. Maher et al., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10797–10801 (2016).
- F. M. Ashcroft, P. Rorsman, *Nat. Rev. Endocrinol.* **9**, 660–669 (2013).
- R. Buzzetti, S. Zampetti, E. Maddaloni, *Nat. Rev. Endocrinol.* **13**, 674–686 (2017).
- D. Jacobs-Tulleneers-Thevissen et al., *Diabetologia* **56**, 1605–1614 (2013).
- T. Desai, L. D. Shea, *Nat. Rev. Drug Discov.* **16**, 338–350 (2017).
- G. W. Zamponi, J. Striessnig, A. Koschak, A. C. Dolphin, *Pharmacol. Rev.* **67**, 821–870 (2015).
- M. D'Arco, A. C. Dolphin, *Sci. Signal.* **5**, pe34 (2012).
- T. Kim et al., *Biochem. Biophys. Res. Commun.* **324**, 401–408 (2004).
- A. Lluçà-Valldeperas et al., *J. Tissue Eng. Regen. Med.* **9**, E76–E83 (2015).
- M. H. Thelen, W. S. Simonides, C. van Hardeveld, *Biochem. J.* **321**, 845–848 (1997).
- Y. Manabe et al., *PLOS ONE* **7**, e52592 (2012).
- J. T. McCluskey et al., *J. Biol. Chem.* **286**, 21982–21992 (2011).
- S. M. Burns et al., *Cell Metab.* **21**, 126–137 (2015).
- J. Gromada, P. Chabosseau, G. A. Rutter, *Nat. Rev. Endocrinol.* **14**, 694–704 (2018).
- H. Yamashita, D. T. Nguyen, E. Chung, *Methods Mol. Biol.* **1098**, 145–151 (2014).
- B. Rattner, in *Host Response to Biomaterials*, S. Badyrak, Ed. (Academic Press, 2015), pp. 37–51.
- S. J. Russell et al., *Lancet Diabetes Endocrinol.* **4**, 233–243 (2016).
- P. A. Hollander et al., *Diabetes Care* **24**, 983–988 (2001).
- S. E. Kahn et al., *J. Clin. Endocrinol. Metab.* **86**, 5824–5829 (2001).
- H. G. Mond, G. Freitag, *Pacing Clin. Electrophysiol.* **37**, 1728–1745 (2014).
- S. J. Russell et al., *N. Engl. J. Med.* **371**, 313–325 (2014).
- G. Schmelzeisen-Redeker et al., *J. Diabetes Sci. Technol.* **9**, 1006–1015 (2015).
- D. B. Keenan, J. J. Mastrototaro, G. Voskanyan, G. M. Steil, *J. Diabetes Sci. Technol.* **3**, 1207–1214 (2009).
- M. Qi, *Adv. Med.* **2014**, 429710 (2014).
- R. R. Henry et al., *Diabetes* **67** (suppl. 1), 138-OR (2018).
- Y. Evron et al., *Sci. Rep.* **8**, 6508 (2018).

ACKNOWLEDGMENTS

We thank the Geneva Islet Transplantation Center and H. Zulewski for human pancreatic islets, which were obtained through the basic research program of the European Consortium for Islet Transplantation (ECIT). We thank S. Bürgel for help in the initial stage of the project, A. Hierlemann for providing pulse generators, M. Folcher for constructive discussions, E. Siringil for support with 3D printing, B. Lang for advice on statistical analysis, A. M. Palma Teixeira and G. Camenisch for preparation of the animal experimentation applications, A. Graff-Meyer and C. Genoud for taking electron microscopy images, and H. Zhao and N. Franko for their help in assembling implants. **Funding:** Supported by a European Research Council (ERC) advanced grant (ElectroGene; grant no. 785800) and in part by the Swiss National Science Foundation (SNF) National Centre of Competence in Research (NCCR) for Molecular Systems Engineering. ECIT was supported by the Juvenile Diabetes Research Foundation (JDRF; grant 31-2008-416). **Author contributions:** K.K. and M.F. designed the project; K.K. and P.S. performed the cell culture experiments and K.K. designed the implants; K.K., S.X., G.C., M.D.H., J.S., and H.Y. performed the animal experiments; P.B. designed the electronic switchboard; K.K., P.S., M.X., and M.F. designed the experiments and analyzed the results; K.K., P.S., M.X., and M.F. wrote the manuscript; and S.X., M.D.H., M.X., and M.F. designed the modified implants and the in vivo “refill” as well as the insulin kinetics experiments. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** The authors declare that all the data supporting the findings of this study are available within the paper and its supplementary materials. Original plasmids are available upon request. All vector information is provided in table S7.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6494/993/suppl/DC1
Materials and Methods
Figs. S1 to S20
Tables S1 to S8
References (59–69)

9 July 2018; resubmitted 11 February 2020

Accepted 3 April 2020

10.1126/science.aau7187

REPORT

SEPARATIONS

Control of zeolite pore interior for chemoselective alkyne/olefin separations

Yuchao Chai^{1*}, Xue Han^{2*}, Weiyao Li², Shanshan Liu¹, Sikai Yao¹, Chong Wang³, Wei Shi⁴, Ivan da-Silva⁵, Pascal Manuel⁵, Yongqiang Cheng⁶, Luke D. Daemen⁶, Anibal J. Ramirez-Cuesta⁶, Chiu C. Tang⁷, Ling Jiang³, Sihai Yang^{2†}, Naijia Guan^{1,4}, Landong Li^{1,4†}

The efficient removal of alkyne impurities for the production of polymer-grade lower olefins remains an important and challenging goal for many industries. We report a strategy to control the pore interior of faujasite (FAU) zeolites by the confinement of isolated open nickel(II) sites in their six-membered rings. Under ambient conditions, Ni@FAU showed remarkable adsorption of alkynes and efficient separations of acetylene/ethylene, propyne/propylene, and butyne/1,3-butadiene mixtures, with unprecedented dynamic separation selectivities of 100, 92, and 83, respectively. In situ neutron diffraction and inelastic neutron scattering revealed that confined nickel(II) sites enabled chemoselective and reversible binding to acetylene through the formation of metastable $[\text{Ni}(\text{II})(\text{C}_2\text{H}_2)_3]$ complexes. Control of the chemistry of pore interiors of easily scalable zeolites has unlocked their potential in challenging industrial separations.

More than 350 million metric tons of lower olefins (ethylene, propylene, and 1,3-butadiene) are produced each year through the steam cracking of hydrocarbons. Separating large quantities of chemical mixtures into purer forms accounts for an enormous amount of global energy consumption (1). To obtain polymer-grade olefins, the by-products of alkynes (acetylene, propyne, and butyne) in the stream must be reduced to <5 parts per million (ppm), because these alkynes irreversibly poison the catalysts for polymerization (2). State-of-the-art techniques to purify olefins are based on the partial hydrogenation of alkynes over supported Pd-catalysts; however, such methods suffer from poor selectivity and high costs (3). Emerging porous sorbents, notably metal-organic frameworks (MOFs), show preferential adsorption of alkynes over olefins, suggesting alternative adsorption-based purification processes for ethylene (4–8) and propylene (9–11).

However, such processes have yet to be commercialized because of the inherently limited

stability and high production costs of MOFs. Additionally, the primary physisorption mechanism in MOFs that drives separations results in a trade-off between adsorption selectivity and capacity.

Zeolites have structural robustness and low-cost production, and they are widely used for industrial separations on the basis of their molecular sieving property (12), but they are not effective for alkyne/olefin separations because these molecules have similar molecular sizes and volatilities (13). Zeolites can act as useful scaffolds to stabilize active metal sites to uncover previously unidentified functions and properties. In this work, we confine isolated Ni(II) sites into faujasite (FAU) zeolite to achieve remarkable adsorption of alkynes from a range of alkyne/olefin mixtures. The strong yet fully reversible binding between alkyne and the open Ni(II) sites results in the formation of metastable $[\text{Ni}(\text{alkyne})_3]$ complexes under dynamic conditions and enables the complete removal of alkynes from olefins (alkynes <1 ppm). The facile production and high stability of Ni@FAU reinforce its potential in the industrial purification of lower olefins.

M@FAU zeolites $[\text{M} = \text{Ni}(\text{II}), \text{Cu}(\text{II}), \text{and Zn}(\text{II})]$ were synthesized from hydrothermal reactions of mixed gels {molar ratio of $\text{SiO}_2:\text{Al}_2\text{O}_3:\text{Na}_2\text{O}:\text{M-TAPTS}:\text{H}_2\text{O} = 7.8:1.0:2.2:0.6:174$; TAPTS = 3-[2-(2-aminoethylamino)ethylamino]propyltrimethoxysilane} and subsequent processing (yield, 77 to 85%). The TAPTS ligand was used to coordinate Ni(II) ions for their inclusion in the zeolite pore structure at locations that can be difficult to access during conventional post-synthesis ion-exchange. Similar approaches based on (3-mercaptopropyl)trimethoxysilane ligand have been reported to introduce metal

sites or clusters into desirable sites of various zeolites, showing excellent catalytic activities (14, 15). Synchrotron x-ray powder diffraction data confirmed that M@FAU zeolites crystallize in the cubic space group, $Fd\bar{3}m$, adopting the FAU-topology, and there is an absence of bulk phase of metal oxides (fig. S1 and table S1).

The homogeneous distribution of transition-metal cations throughout the M@FAU crystals (3 to 5 μm) were confirmed by electron microscopy (figs. S2 and S3). Under ambient conditions, the pores of M@FAU samples were filled with water molecules that could be removed completely by heating to 623 K (fig. S4). The divalent oxidation state of the confined metal ions was confirmed by x-ray photoelectron spectroscopy (fig. S5), and the primary location of confined Ni(II) sites in FAU zeolite was studied using density functional theory (DFT) calculations (fig. S6) and in situ neutron powder diffraction (NPD) studies.

The adsorption capacities of desolvated M@FAU and the parent Na-FAU were first evaluated by measuring the adsorption isotherms of C_2H_2 and C_2H_4 (Fig. 1A and figs. S7 and S8). At 1 bar and 298 K, Ni@FAU, $[\text{Ni}_{12}\text{Na}_{20}(\text{Al}_{44}\text{Si}_{148}\text{O}_{384})]$, showed lower overall uptakes of C_2H_2 [3.48 mmol/g; equivalent to $\sim 3.7 \text{ C}_2\text{H}_2$ per Ni(II) site] and C_2H_4 (2.36 mmol/g) compared with that of Na-FAU (5.15 and 4.06 mmol/g, respectively), because of the moderate decrease of Brunauer-Emmett-Teller surface areas on incorporation of Ni(II) sites (from 710 to 531 m^2/g ; fig. S9). Notably, 2.0 mmol/g of C_2H_2 uptake was recorded in Ni@FAU at 0.02 bar, a pressure that is relevant to the partial pressure of C_2H_2 impurity in industrial C_2H_4 streams. The steep C_2H_2 uptake of Ni@FAU at low pressure was consistent with its higher heat of adsorption (48.6 kJ/mol) than that of Na-FAU (21.7 kJ/mol) and of C_2H_4 uptake in Ni@FAU (25.8 kJ/mol), which were determined by differential scanning calorimetry (figs. S10 to S21, table S2, and supplementary text).

Temperature-programmed desorption (TPD) profiles of C_2H_2 - and C_2H_4 -loaded Ni@FAU revealed higher adsorption uptake of C_2H_2 (1.75 mmol/g) than of C_2H_4 (0.51 mmol/g) and showed desorption peaks centered at 363 and 329 K, respectively. When coadsorbed with an equimolar mixture of $\text{C}_2\text{H}_2/\text{C}_2\text{H}_4$, the TPD profile resembled that of C_2H_2 -loaded Ni@FAU, and little desorption of C_2H_4 was observed. Moreover, the pre-adsorbed C_2H_4 molecules in Ni@FAU could be readily displaced by C_2H_2 under dynamic conditions (Fig. 1B). The presence of strongly bound C_2H_2 molecules in Ni@FAU is also confirmed by in situ Fourier transform infrared (FTIR) studies at 298 K (Fig. 1C). On adsorption in Ni@FAU, $\nu_{\text{as}}(\text{CH})$ and $\nu_{\text{s}}(\text{CH})$ bands of C_2H_2 red-shifted to 2925 and 3010 cm^{-1} , respectively, compared with gaseous or physisorbed C_2H_2 (3100 to 3400 cm^{-1}) (16, 17), whereas no shift was observed for adsorbed C_2H_4 molecules.

¹School of Materials Science and Engineering and National Institute for Advanced Materials, Nankai University, Tianjin 300350, China. ²Department of Chemistry, The University of Manchester, Manchester, M13 9PL, UK. ³State Key Laboratory of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China. ⁴Key Laboratory of Advanced Energy Materials Chemistry of Ministry of Education, Nankai University, Tianjin 300071, China. ⁵ISIS Facility, Science and Technology Facilities Council (STFC), Rutherford Appleton Laboratory, Chilton, Oxfordshire, OX11 0QX, UK. ⁶Neutron Scattering Division, Neutron Sciences Directorate, Oak Ridge National Laboratory (ORNL), Oak Ridge, TN 37831, USA. ⁷Diamond Light Source, Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0DE, UK.

*These authors contributed equally to this work.

†Corresponding author. Email: sihai.yang@manchester.ac.uk (S.Y.); lild@nankai.edu.cn (L.L.)

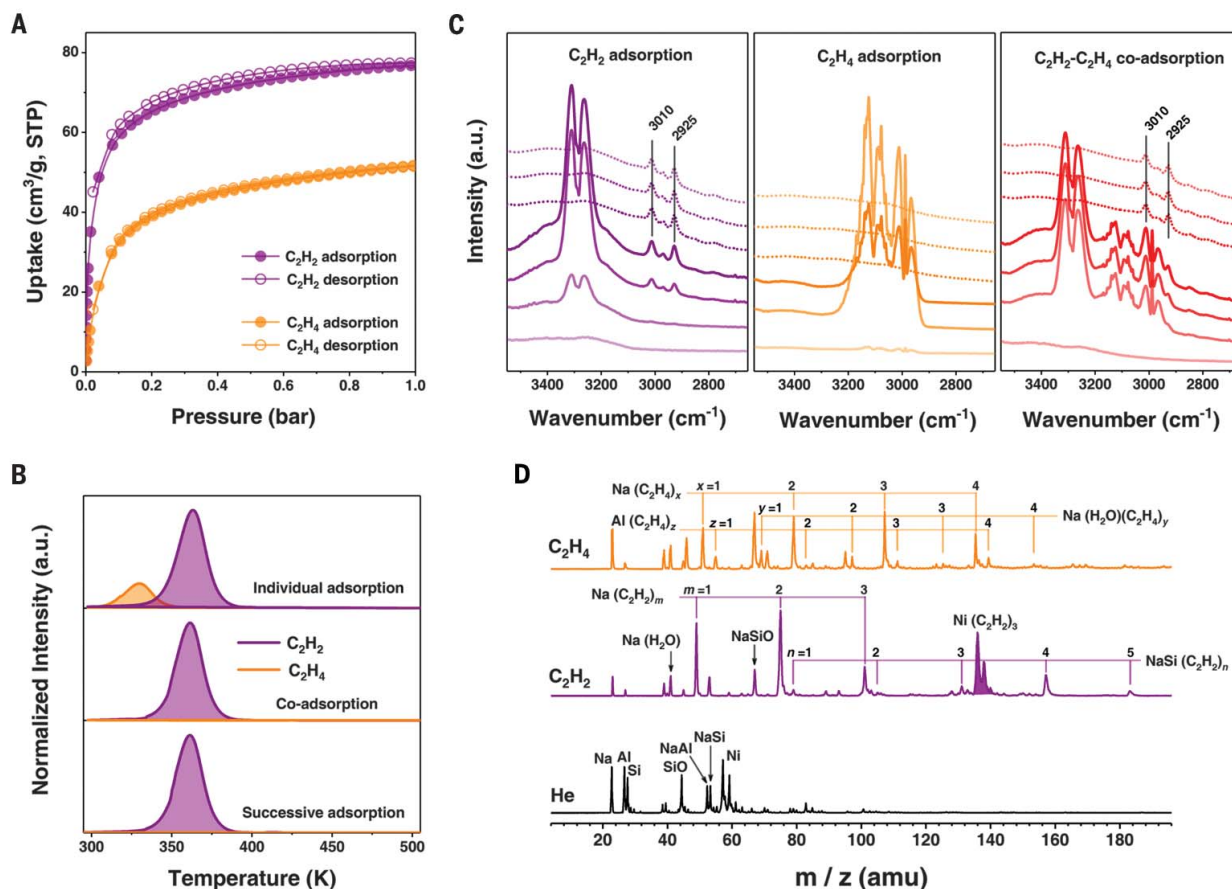


Fig. 1. Adsorption data of C₂H₂ and C₂H₄ for Ni@FAU. (A) Adsorption isotherms of C₂H₂ and C₂H₄ for Ni@FAU at 298 K. STP, standard temperature and pressure. (B) TPD profiles of C₂H₂- and C₂H₄-adsorbed Ni@FAU after their individual adsorption, coadsorption, and successive adsorption (first C₂H₄ and then switched to C₂H₂) at 298 K. a.u., arbitrary units.

(C) In situ FTIR spectra of Ni@FAU on adsorption of C₂H₂ and C₂H₄ followed by He purging (dotted lines) at 298 K. (D) Mass spectra of species produced by pulsed laser vaporization of the Ni@FAU target in the presence of carrier gas He, C₂H₂ (2%)/He, and C₂H₄ (2%)/He. m/z, mass/charge ratio; amu, atomic mass unit.

Infrared bands of bound C₂H₂ molecules were observed in the equimolar C₂H₂/C₂H₄ co-adsorbed Ni@FAU, demonstrating the selective uptake of C₂H₂ under competitive adsorption.

The adsorption species in C₂H₂- and C₂H₄-loaded Ni@FAU were identified by mass spectrometry (18). The pulsed laser vaporization of Ni@FAU target in He produced a series of fragments, and Ni-containing fragments can be distinguished by the characteristic isotope ratio of nickel (⁵⁸Ni:⁶⁰Ni = 68%:26%). On the basis of the control experiment of metallic Ni and Na-FAU (figs. S22 and S23), fragments corresponding to Ni(C₂H₂)₃ species were identified as a key species in C₂H₂-adsorbed Ni@FAU (Fig. 1D), whereas no Ni(C₂H₄)_n (*n* = 1 to 4) species were observed for C₂H₄-adsorbed Ni@FAU. These results demonstrated the highly selective adsorption of C₂H₂ in Ni@FAU and its high capability to remove trace C₂H₂ from the C₂H₄ stream.

The ability of M@FAU (M = Ni, Cu, and Zn) to separate C₂H₂/C₂H₄ mixtures under dynamic conditions was evaluated by breakthrough experiments. All of the samples showed sufficient dynamic adsorption of C₂H₂ and could

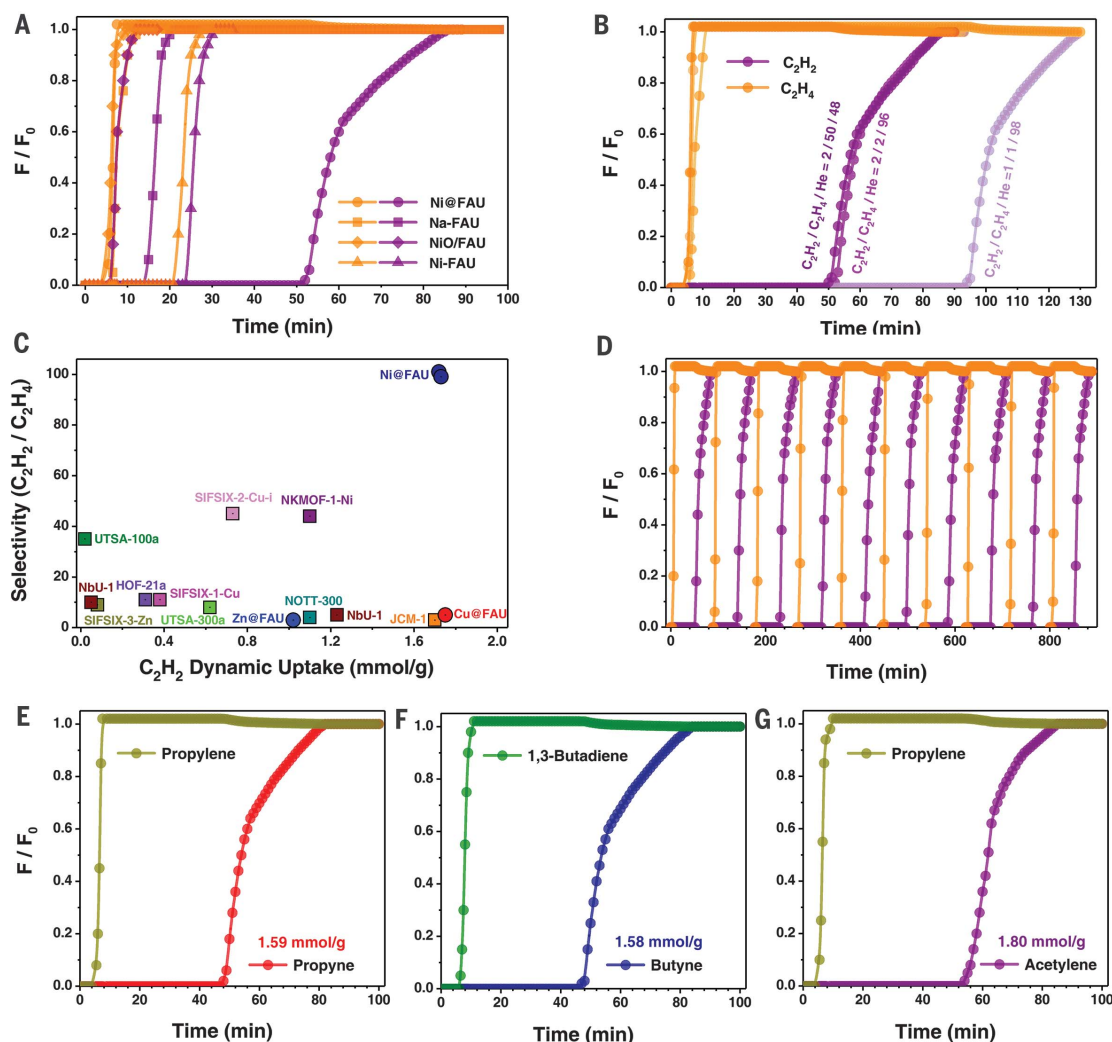
produce ultrapure C₂H₄ (C₂H₂ <1 ppm) streams at the outlet of the fixed beds at 298 K (Fig. 2A and figs. S24 and S25). The dynamic C₂H₂ uptakes calculated from column breakthrough curves (2% C₂H₂:2% C₂H₄) were 0.91, 1.26, and 1.72 mmol/g for Zn@FAU, Cu@FAU, and Ni@FAU, respectively. The dynamic C₂H₂ uptake of Ni@FAU remained the same when the C₂H₄ concentration was increased to 50% or the C₂H₂ concentration decreased to 1% (Fig. 2B). This capability is distinct to the porous sorbents functioning solely through physisorption (6), which show rapid reductions of dynamic uptakes with decreasing gas concentrations.

Furthermore, this dynamic uptake compared favorably to the leading MOFs—for example, 1.70 mmol/g of JCM-1 (19), 1.18 mmol/g of UTSA-200a (20), 1.23 mmol/g of NbU (21), and 0.73 mmol/g of SIFSIX-2-Cu-i (4). Additionally, Ni@FAU showed a higher overall acetylene productivity (116.8 mmol/g) compared with that of leading MOFs, such as NKMOF-1-Ni (96.0 mmol/g) (8), UTSA-200a (85.7 mmol/g) (20), and SIFSIX-2-Cu-i (53.3 mmol/g) (4). Also, Ni@FAU exhibited a very low dynamic uptake

of C₂H₄ (0.02 mmol/g) under the same conditions, and a remarkable C₂H₂/C₂H₄ dynamic selectivity of ~100 was achieved, compared with reported state-of-the-art sorbents (table S3). A comparison of experimentally determined C₂H₂/C₂H₄ dynamic selectivity against the dynamic uptake of C₂H₂ demonstrated the remarkable performance of Ni@FAU for C₂H₂/C₂H₄ separation (Fig. 2C).

We observed similar separation capability with Ni@FAU after increasing the column temperature from 298 to 308 K (fig. S26) or adding CO₂ or H₂O into the gas stream (fig. S27), whereas an ~25% decrease in the adsorption capability of C₂H₂ was observed for a Ni@FAU sample that has been preadsorbed with H₂O (fig. S28). The performance of Ni@FAU was further evaluated at 5 bar, where excellent C₂H₂/C₂H₄ separations were observed and the dynamic uptakes of C₂H₂ and C₂H₄ increased to 2.25 and 0.39 mmol/g, respectively (fig. S29). Moreover, Ni@FAU could be used to separate mixtures of propyne/propylene, butyne/1,3-butadiene, and acetylene/propylene with high dynamic alkyne uptakes of 1.58 to 1.80 mmol/g

Fig. 2. Column breakthrough studies for alkyne/olefin separations. (A) Column breakthrough curves for a C_2H_2/C_2H_4 (2%/2%) mixture using various zeolite samples at 298 K. C_2H_2 and C_2H_4 are shown in purple and orange, respectively. F, flow rate; F_0 , initial flow rate. (B) Effects of feed gas composition on C_2H_2/C_2H_4 separation over Ni@FAU at 298 K. (C) Plot of C_2H_2/C_2H_4 dynamic selectivity against C_2H_2 dynamic uptake under ambient conditions with state-of-the-art sorbent materials. (D) View of recyclability of Ni@FAU for the separation of C_2H_2/C_2H_4 (2%/2%) at 298 K. Sample regeneration was achieved by treatment in He at 423 K for 30 min. (E to G) Column breakthrough curves for propyne/propylene (2%/2%) (E), butyne/1,3-butadiene (2%/2%) (F), and acetylene/propylene (2%/2%) (G) over fixed beds packed with Ni@FAU at 298 K. Total gas flow, 6.0 mL/min; sample weight, 0.2 g.



and notable breakthrough selectivities of 92, 83, and 109, respectively, under ambient conditions (Fig. 2, E to G).

These results signal the potential of Ni@FAU for the adsorptive removal of alkynes from industrial olefin streams. Practical sorbents must be recyclable. After 10 cycles of C_2H_2/C_2H_4 separations with Ni@FAU, we observed no decline in the retention time and full sorbent regeneration at 373 K between each cycle (Fig. 2D). In contrast, Cu@FAU exhibited poor reversibility for C_2H_2/C_2H_4 separations because of the formation of oligomers on Cu(II) sites blocking the pores (fig. S30). To understand the role of Ni(II) in Ni@FAU, we introduced the metal ions into the FAU zeolites using different methods, such as ion-exchange (denoted as Ni-FAU) or wet impregnation (denoted as NiO/FAU) (figs. S31 and S32). These two samples exhibited very poor separation of C_2H_2/C_2H_4 (Fig. 2A), which suggests that the excellent performance of Ni@FAU originates from its binding environment of the confined Ni(II) sites within the pores.

In situ NPD studies enabled identification of the locations of the confined Ni(II) sites and the adsorbed gas (C_2D_2 , C_2D_4 , C_3D_4 , and C_3D_6) molecules within Ni@FAU (figs. S33 to S38 and table S4). Fourier difference map analysis of the desolvated Ni@FAU [$Ni_{12}Na_{20}(Al_{44}Si_{148}O_{384})$] confirmed the structural integrity and the absence of residual nuclear density in the supercage. In-depth analysis revealed apparent residual nuclear density near the six-membered ring of the sodalite cage, which we assigned as Ni(II) ions stabilized by framework oxygen centers (fig. S39). This assignment is consistent with DFT calculations (fig. S6). By contrast, Ni(II) sites within Ni-FAU were primarily hexagonal prism sites, which were sterically hindered by the highly confined void (diameter of ~ 2.5 Å) for gas binding. This finding was consistent with its poor separation performance (fig. S40 and supplementary text).

Upon gas loading, variations in Bragg peak intensities were observed, and the binding domains of gases were successfully interpreted by Fourier difference map analysis (Fig. 3, A to E)

and Rietveld refinements. At low loading in [$Ni_{12}Na_{20}(Al_{44}Si_{148}O_{384})$] $\cdot(C_2D_2)_{12}$ (equivalent to ~ 1 mmol/g C_2D_2 uptake), all adsorbed C_2D_2 molecules were located at a single site, which is distributed over six equivalent positions in the supercage and exhibited a side-on interaction to the Ni(II) sites [$C \cdots Ni = 3.87$ to 4.08 Å, $\angle C \equiv C \cdots Ni = 91.1^\circ$], which suggests the binding interaction between the $C \equiv C$ bond and Ni(II) centers (fig. S41). Additionally, supplementary hydrogen bonds between $D_{C_2D_2}$ and framework oxygen were observed [$D \cdots O = 2.99$ Å, $\angle C-D \cdots O = 161^\circ$] (Fig. 3B). A similar host-guest binding mechanism [$C \cdots Ni = 3.83$ to 4.05 Å, $\angle C \equiv C \cdots Ni = 91.9^\circ$] was observed in [$Ni_{12}Na_{20}(Al_{44}Si_{148}O_{384})$] $\cdot(C_2D_2)_{26}$ (equivalent to ~ 2.3 mmol/g C_2D_2 uptake) with weaker hydrogen bonds to the framework oxygen [$D \cdots O = 3.27$ Å, $\angle C-D \cdots O = 151^\circ$]. The overall binding geometry is in marked agreement with cation-acetylene π complexation (13, 22). X-ray absorption near-edge structure (XANES) analysis confirmed the retention of a divalent oxidation state of Ni(II) sites on

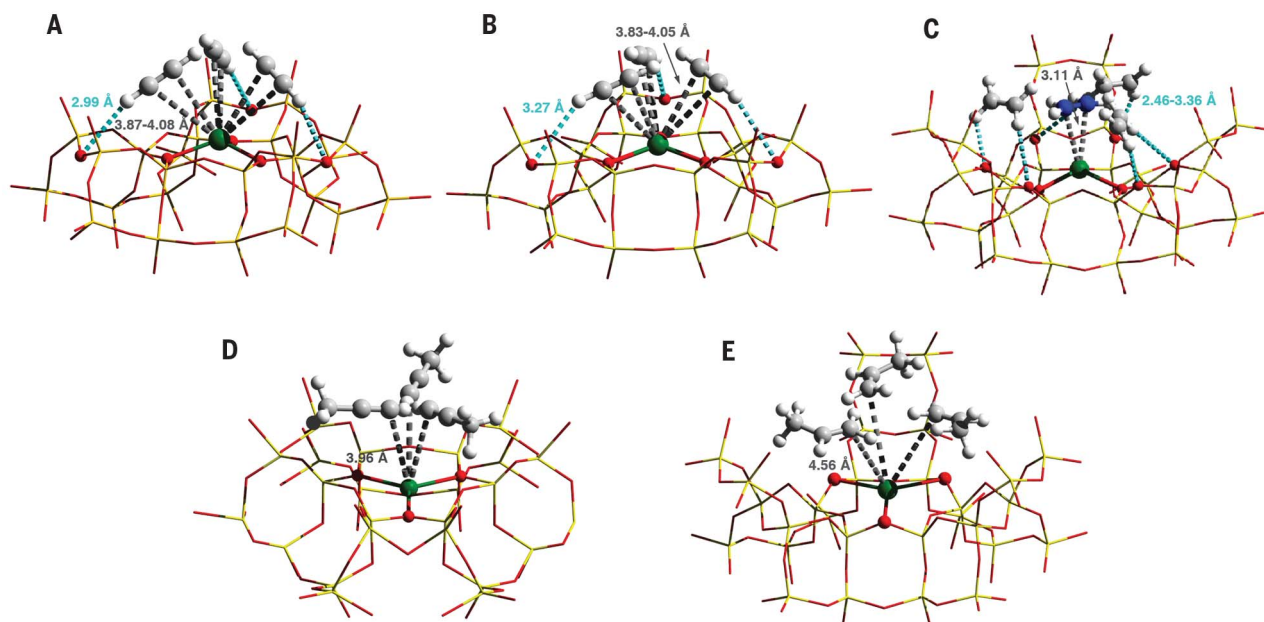


Fig. 3. Views of crystal structures for the Ni@FAU zeolite as a function of gas loading. All structures were derived from Rietveld refinements of NPD data at 7 K [Si and Al: yellow; O: red; Ni: green; C: gray; D: white; C_2D_4 (1) is highlighted in blue for clarity]. The host-guest interactions are highlighted by dashed lines, and the estimated

standard deviation values for binding distances are typically within 0.02 to 0.08 Å. Views are of binding sites for adsorbed gas molecules in $[Ni_{12}Na_{20}(Al_{44}Si_{148}O_{384})] \cdot (C_2D_2)_{12}$ (A), $[Ni_{12}Na_{20}(Al_{44}Si_{148}O_{384})] \cdot (C_2D_2)_{26}$ (B), $[Ni_{12}Na_{20}(Al_{44}Si_{148}O_{384})] \cdot (C_2D_4)_{17}$ (C), $[Ni_{12}Na_{20}(Al_{44}Si_{148}O_{384})] \cdot (C_3D_4)_{26}$ (D), and $[Ni_{12}Na_{20}(Al_{44}Si_{148}O_{384})] \cdot (C_3D_6)_{26}$ (E).

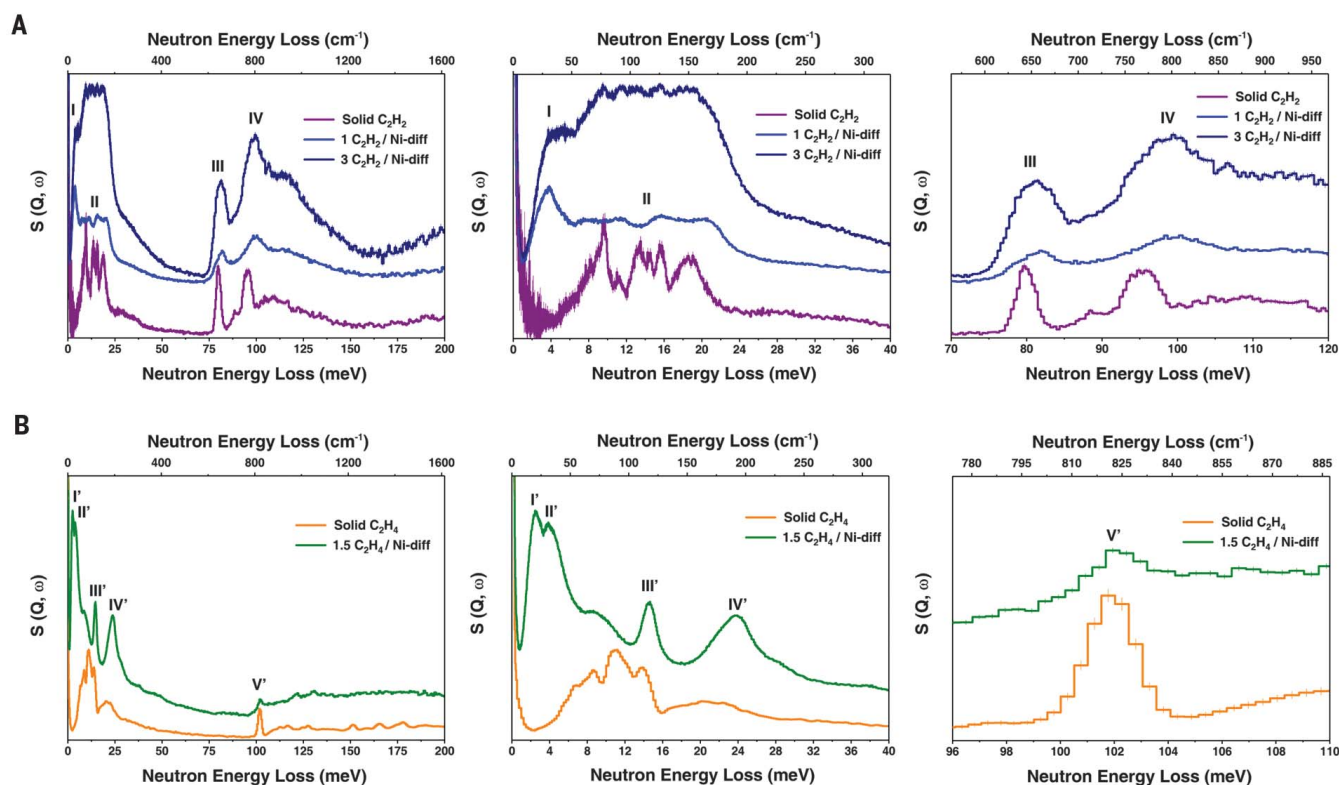


Fig. 4. INS spectra for Ni@FAU as a function of gas loading. (A) Comparison of INS spectra of C_2H_2 -loaded Ni@FAU and that of solid C_2H_2 . (B) Comparison of INS spectra of C_2H_4 -loaded Ni@FAU and that of solid C_2H_4 . Enlarged details show the translational or librational and the internal vibrational modes of

adsorbed C_2H_2 and C_2H_4 molecules. Difference spectra were produced by removing signals of the bare zeolite and sample holder. Raw spectra are provided in the supplementary materials. Peaks are labeled with Roman numerals. S, dynamic structure factor; Q, momentum transfer; ω , frequency change.

acetylene binding (fig. S42), and no elongation of C–C distance in bound C_2D_2 was found in the NPD analysis. These results were consistent with the selective yet reversible sorption.

However, $[Ni_{12}Na_{20}(Al_{44}Si_{148}O_{384})] \cdot (C_2D_4)_{17}$ (equivalent to ~ 1.5 mmol/g C_2D_4 uptake) exhibited a different binding geometry with two distinct sites (**1** and **2**) in the supercage (fig. S41). C_2D_4 (**1**) molecules (accounting for $\sim 40\%$ of adsorbed C_2D_4) interact with Ni(II) [$C \cdots Ni = 3.11 \text{ \AA}$, $\angle C = C \cdots Ni = 77.8^\circ$] in a similar side-on manner to that of C_2D_2 , whereas C_2D_4 (**2**) (accounting for $\sim 60\%$ of adsorbed C_2D_4) showed no interaction with Ni(II) but did show multiple hydrogen bonds to the framework oxygen [$D \cdots O = 2.46$ to $3.36 \text{ \AA}$, $\angle C - D \cdots O = 124$ to 178°]. These NPD studies revealed the explicit difference between C_2D_2 and C_2D_4 upon adsorption in Ni@FAU. Additional π electrons in C_2D_2 and its linear geometry (and thus low spatial hindrance) enabled the formation of metastable $[Ni(II)(C_2D_2)_3]$ complexes in the supercage of FAU, which is fully consistent with the mass spectrometry results. In contrast, the Ni(II) sites were heavily blocked by the bulky C_2D_4 by the formation of a dynamic 1:1 adduct, which led to most of the adsorbed C_2D_4 molecules being stabilized through weak hydrogen bonding and intermolecular guest-guest interactions. Similar host-guest binding interactions were observed in the structure models of $[Ni_{12}Na_{20}(Al_{44}Si_{148}O_{384})] \cdot (C_3D_4)_{20}$ and $[Ni_{12}Na_{20}(Al_{44}Si_{148}O_{384})] \cdot (C_3D_6)_{26}$ with $C \cdots Ni$ distances of 3.96 and 4.56 $\text{\AA}$, respectively (only one binding site observed in each model; Fig. 3, D and E). Thus, the distinct nature of sorbent-gas interaction construed the high selectivity of Ni@FAU toward alkyne adsorption.

To visualize the binding dynamics of adsorbed C_2H_2 and C_2H_4 molecules, inelastic neutron scattering (INS) studies were conducted with Ni@FAU as a function of gas loading at 5 K (figs. S43 to S48). Compared with the INS spectra of solid C_2H_2 and C_2H_4 , the difference INS spectra (i.e., signals of adsorbed gas molecules) had differences in the low-energy region (<30 meV) that were correlated to the translation and libration modes of molecular vibration. In the solid state, molecules interact with adjacent ones in all three dimensions (figs. S49 to S54 and table S5), which results in coupling and dispersion of the modes (23). When adsorbed onto Ni(II) sites, gas molecules become isolated and restricted in an anisotropic environment, which results in distinct INS features. The peak frequencies and (anisotropic) amplitude were directly dictated by the Ni(II)-gas interactions and the local environment.

We assigned the sharp and intense INS peaks at 3.8 meV for bound C_2H_2 (peak I) and 2.5 to 3.8 meV for adsorbed C_2H_4 (peaks I' and II') to the motion of gas molecules within the

plane perpendicular to the Ni(II)-gas axis, because these vibrational modes are the least hindered and had the lowest frequencies and largest displacement. These INS spectra were in marked agreement with the binding sites elucidated by NPD, in terms of peak I correlating to the sole binding site of C_2H_2 , whereas peaks I' and II' resulting from the two C_2H_4 sites with the former being the major adsorption site. Peak I for bound C_2H_2 occurred at a higher average energy, which confirmed that the interaction between Ni@FAU and C_2H_2 was stronger than that of C_2H_4 .

Additionally, adsorbed gas molecules may rotate or twist around the axis perpendicular to the Ni(II)-gas axis and move toward or away from the Ni(II) sites, contributing to the peaks at 10 to 15 meV and at 15 to 30 meV, respectively. Moreover, the trans- and cis-C–H bending modes for bound C_2H_2 at 81.5 and 99 meV, respectively, were blue-shifted compared with those of solid C_2H_2 (80 and 95 meV, respectively; Fig. 4A), which indicates that these internal modes were strongly hindered upon binding on the Ni(II) sites. In contrast, no apparent shift was observed for the in-plane C–H rocking mode of C_2H_4 at 102 meV upon adsorption (Fig. 4B). Overall, the INS study showed marked agreement with NPD, adsorption, and breakthrough results, and it identified the crucial role of confined Ni(II) sites on the chemoselective binding of C_2H_2 in Ni@FAU.

Solid-sorbent-based techniques hold increasing promise to improve the operational efficiency of existing separation processes in petrochemical industries, and the separation of alkyne impurities from olefins can only be realized by exploiting the differences in their properties, such as dimensions (24, 25), shapes (26), conformation (27), polarisabilities (4), coordination abilities (28), binding affinity (29), and the geometry-matching with the sorbent pores (30). Zeolites with well-defined channels have been considered to be viable candidates for gas separation for decades, primarily on the basis of their molecular sieving property (12). By confining atomically dispersed Ni(II) sites in the FAU zeolite channels, the discrimination between alkyne and olefin binding was amplified in Ni@FAU, which enabled the production of polymer-grade olefins under conditions relevant to practical processes. Combining its facile synthesis at large scale and its notable stability, the Ni@FAU sorbent offers a potential practical solution to the challenging alkyne/olefin separations.

REFERENCES AND NOTES

- D. S. Sholl, R. P. Lively, *Nature* **532**, 435–437 (2016).
- H. Zimmermann, R. Walz, in *Ullmann's Encyclopedia of Industrial Chemistry* (Wiley, 2000).
- A. Borodziński, G. C. Bond, *Catal. Rev., Sci. Eng.* **48**, 91–144 (2006).
- X. Cui et al., *Science* **353**, 141–144 (2016).
- S.-C. Xiang et al., *Nat. Commun.* **2**, 204 (2011).

- S. Yang et al., *Nat. Chem.* **7**, 121–129 (2015).
- T. L. Hu et al., *Nat. Commun.* **6**, 7328–7335 (2015).
- Y.-L. Peng et al., *Angew. Chem. Int. Ed.* **57**, 10971–10975 (2018).
- L. Li et al., *J. Am. Chem. Soc.* **139**, 7733–7736 (2017).
- L. Li et al., *Angew. Chem. Int. Ed.* **57**, 15183–15188 (2018).
- L. Yang et al., *Adv. Mater.* **30**, 1705374 (2018).
- R. T. Yang, *Gas Separation by Adsorption Processes* (Imperial College Press, 1997).
- S. Patai, *The Chemistry of the Carbon-Carbon Triple Bond* (Wiley, 1978).
- M. Choi, Z. Wu, E. Iglesia, *J. Am. Chem. Soc.* **132**, 9129–9137 (2010).
- S.-W. Choi et al., *J. Catal.* **345**, 113–123 (2017).
- J. Erkelens, W. J. Wösten, *J. Catal.* **54**, 143–154 (1978).
- M. P. Lapinski, J. G. Ekerdt, *J. Phys. Chem.* **94**, 4599–4610 (1990).
- P. Wang et al., *Angew. Chem. Int. Ed.* **56**, 8716–8720 (2017).
- J. Lee et al., *Angew. Chem. Int. Ed.* **57**, 7869–7873 (2018).
- B. Li et al., *Adv. Mater.* **29**, 1704210 (2017).
- J. Li et al., *J. Am. Chem. Soc.* **141**, 3807–3811 (2019).
- J. H. Nelson, H. B. Jonassen, *Coord. Chem. Rev.* **6**, 27–63 (1971).
- P. C. H. Mitchell, S. F. Parker, A. J. Ramirez-Cuesta, J. Tomkinson, *Vibrational Spectroscopy with Neutrons with Applications in Chemistry, Biology, Material Science and Catalysis* (World Scientific Publishing, 2005).
- A. Cadiou, K. Adil, P. M. Bhatt, Y. Belmabkhout, M. Eddaoudi, *Science* **353**, 137–140 (2016).
- P. J. Bereciartua et al., *Science* **358**, 1068–1071 (2017).
- Z. R. Herm et al., *Science* **340**, 960–964 (2013).
- P.-Q. Liao, N.-Y. Huang, W.-X. Zhang, J.-P. Zhang, X.-M. Chen, *Science* **356**, 1193–1196 (2017).
- E. D. Bloch et al., *Science* **335**, 1606–1610 (2012).
- K. Lee, J. D. Howe, L.-C. Lin, B. Smit, J. B. Neaton, *Chem. Mater.* **27**, 668–678 (2015).
- S. M. Kuznicki et al., *Nature* **412**, 720–724 (2001).

ACKNOWLEDGMENTS

We thank C. Wang from China Petrochemical Corporation (SINOPEC) for helpful discussions. **Funding:** Supported by the National Natural Science Fund of China (21722303, 21421001, and 21688102), the Municipal Natural Science Fund of Tianjin (18JCJC47400 and 18JCZDJC37400), the 111 Project (B12015 and B18030), the U.K. Engineering and Physical Sciences Research Council (EP/P011632/1), and the University of Manchester and Newton Advanced Fellowship by the Royal Society (to W.S. and S.Yan.). Beamlines WISH and I11 were accessed through the STFC ISIS Facility and Diamond Light Source, respectively. INS experiments were conducted at the VISION beamline at ORNL's Spallation Neutron Source, which is supported by the Scientific User Facilities Division, Office of Basic Energy Sciences (BES), U.S. Department of Energy (DOE), under contract no. DE-AC0500OR22725 with UT Battelle, LLC. The computing resources were made available through the VirtuES and the ICEMAN projects, funded by the Laboratory Directed Research and Development program at ORNL. **Author contributions:** Y.Cha. conducted zeolite syntheses and performance tests. S.L. and S.Yao participated in sample characterization. C.W. and L.J. performed mass spectrometry analyses. W.S. and N.G. analyzed the data and provided helpful discussions. X.H., W.L., L.-D.-S., P.M., Y.Che., L.D.-S., A.J.R.-C., and C.C.T. collected and analyzed synchrotron x-ray diffraction, neutron diffraction, and INS data. S.Yan. and L.L. directed and supervised the project. Y.Cha., X.H., S.Yan., and L.L. prepared the manuscript. **Competing interests:** The authors from Nankai University are inventors of a patent (CN202010177858.0) on materials reported in this work. **Data and materials availability:** Crystallographic data for the structures reported in this article have been deposited at the Cambridge Crystallographic Data Centre under deposition numbers CCDC 1939577 to 1939585. Copies of the data can be obtained free of charge from www.ccdc.cam.ac.uk/structures/. Other data needed to evaluate the conclusions in the Report are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6494/1002/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S55
Tables S1 to S5
References (31–55)

23 July 2019; resubmitted 16 January 2020
Accepted 10 April 2020
10.1126/science.ayy8447

ORGANIC CHEMISTRY

Total synthesis of bryostatin 3

Barry M. Trost*, Youliang Wang, Andreas K. Buckl, Zhongxing Huang,
Minh H. Nguyen, Olesya Kuzmina

Bryostatins are a family of 21 complex marine natural products with a wide range of potent biological activities. Among all the 21 bryostatins, bryostatin 3 is structurally the most complex. Whereas nine total syntheses of bryostatins have been achieved to date, bryostatin 3 has only been targeted once and required the highest number of steps to synthesize (43 steps in the longest linear sequence and 88 total steps). Here, we report a concise total synthesis of bryostatin 3 using 22 steps in the longest linear sequence and 31 total steps through a highly convergent synthetic plan by the use of highly atom-economical and chemoselective transformations in which alkynes played a major role in reducing step count.

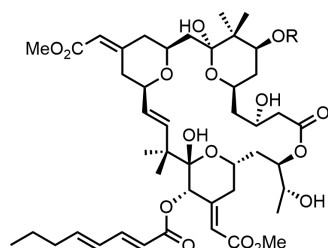
The bryostatins, first isolated by Pettit *et al.* (1, 2) from the marine bryozoan *Bugula neritina*, are a family of 21 macrolides (3–6) with potent antineoplastic (7, 8), immunopotentiating (9), synaptogenesis-inducing (10), and latent HIV-modulating (11) activity. Beneficial effects as a post-stroke treatment (12) and for restoring the blood-brain barrier after traumatic blast injuries (13) have also been demonstrated. Although the exact mechanism of action remains an ongoing area of research (14), it has become clear that bryostatins act as agonists of protein kinase C (PKC), with low nanomolar affinities for their target. Their pharmacological potential together with their intriguingly complex structures have attracted the attention of numerous synthetic organic chemists over the past 30 years. To date, nine completed total syntheses have been reported (Fig. 1): bryostatin 1 (Keck, 2011; Wender, 2017) (15, 16), bryostatin 2 (Evans, 1999) (17), bryostatin 3 (Yamamura, 2000) (18), bryostatin 7 (Masamune, 1990; Krische, 2011) (19, 20), bryostatin 8 (Song, 2018) (21), bryostatin 9 (Wender, 2011) (22), and bryostatin 16 (Trost, 2008) (23). In addition, Hale has developed a formal synthesis of bryostatin 7 (24–27), and other groups also have made important contributions to this area (28–32). All of these synthetic studies serve as guidelines toward the ultimate goal of a concise route to bryostatins that is practical and flexible, addressing the supply problem and allowing for ready access to analogs for structure-activity-relationship (SAR) studies.

Topologically, all bryostatins share a 26-membered lactone and three highly functionalized tetrahydropyrans integrated in the macrocycle. Bryostatin 3, 19, and 20 are rather exceptional because of an additional butenolide unit directly fused to the macrocycle (Fig. 1, bryostatin 3, for example). This distinguishing butenolide unit is found in

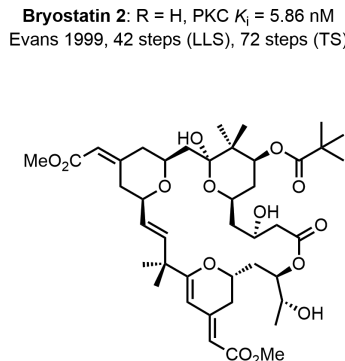
the region of the molecule crucial for binding to its biological target, PKC (33). Although changes in this recognition domain substantially attenuated binding affinities in SAR studies, bryostatin 3 retained low nanomolar affinity for PKC [inhibition constant (K_i) = 2.75 nM] (33), comparable with that of bryostatin 1 (K_i = 1.35 nM). From a synthetic perspective, the butenolide unit renders such bryostatins, especially bryostatin 3, structurally more complex and thus more challenging to synthesize. Among all the completed syntheses (15–23), only one

targeted the butenolide-containing bryostatin: the Yamamura group's synthesis of bryostatin 3 (18). Not surprisingly, both the longest linear step count (43) and the total step count (88) are highest of all the bryostatin syntheses. With the goal of developing practical and flexible approaches to the bryostatin family by using atom-economical and selective reactions, we report a concise total synthesis of bryostatin 3 by using only 22 steps in the longest linear sequence and 31 total steps.

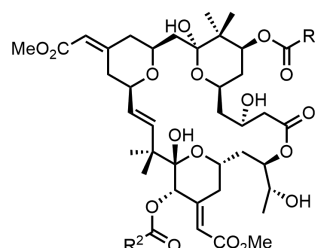
Retrosynthetically, we proposed that bryostatin 3 could be synthesized from the macrocyclic intermediate **1** through a late-stage oxidative functionalization of the dihydropyran ring C and a palladium-catalyzed carbonylative esterification of the stereodefined vinyl bromide (Fig. 2). Intermediate **1** could then be disassembled into intermediate **2** and fragment **3** on the basis of three key transformations: the palladium-catalyzed alkyne-alkyne coupling reaction to construct the C20–C21 bond, the gold-catalyzed 6-*endo-dig* cyclization to generate the dihydropyran ring C, and the Yamaguchi macrolactonization to form the macrocycle from the C1 carboxylic acid and the C25 alcohol. Furthermore, intermediate **2** could be disconnected into fragments **4** and **5** based on another two key



Bryostatin 1: R = Ac, PKC K_i = 1.35 nM
Keck 2011, 31 steps (LLS), 58 steps (TS)
Wender 2017, 19 steps (LLS), 29 steps (TS)



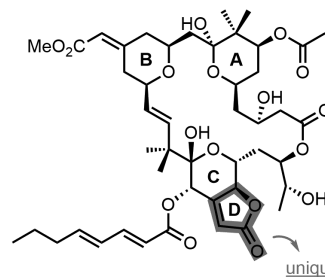
Bryostatin 2: R = H, PKC K_i = 5.86 nM
Evans 1999, 42 steps (LLS), 72 steps (TS)



Bryostatin 7: R¹ = Me, R² = Me, PKC K_i = 0.84 nM
Masamune 1990, 41 steps (LLS), 79 steps (TS)
Krische 2011, 20 steps (LLS), 36 steps (TS)

Bryostatin 8: R¹ = ⁿPr, R² = ⁿPr, PKC K_i = 1.72 nM
Song 2018, 29 steps (LLS), 51 steps (TS)

Bryostatin 9: R¹ = Me, R² = ⁿPr, PKC K_i = 1.31 nM
Wender 2011, 25 steps (LLS), 43 steps (TS)



Bryostatin 3: PKC K_i = 2.75 nM
Yamamura 2000, 43 steps (LLS), 88 steps (TS)
This work, 22 steps (LLS), 31 steps (TS)

Fig. 1. Prior total syntheses of bryostatins.

Department of Chemistry, Stanford University, Stanford, CA 94305-5080, USA.

*Corresponding author. Email: bmtrost@stanford.edu

reactions: the ruthenium-catalyzed alkene-alkyne coupling reaction to construct the C12–C13 bond and the intramolecular Michael addition reaction between the pendant C15 hydroxy group and the in situ generated enone to furnish the tetrahydropyran ring B. With this synthetic plan, bryostatin 3 was envisioned to be accessed from three basic fragments (**3**, **4**, and **5**) with comparable complexity in a highly convergent fashion.

The synthesis of fragment **3** commenced with the Sharpless asymmetric dihydroxylation of 3-pentenitrile to furnish diol **7** (Fig. 3A) (20). Subsequent protection of the diol as cyclopentylidene acetal **9** and partial reduction of the nitrile by diisobutylaluminum hydride (DIBAL-H) afforded aldehyde **10**. Use of the Stork-modified Wittig reaction (34) produced (*Z*)-vinyl iodide **11** in good yield as a single geometric isomer, which was then cross-coupled with methyl propiolate to generate alkynoate **12** with no loss of the *Z*-olefin stereochemistry (35). Another asymmetric dihydroxylation of enyne **12** furnished fragment **3** in excellent yield and moderate diastereoselectivity. Turning our attention toward diyne fragment **4**, we developed a three-step synthesis from 3-methyl butyne **13**. Double lithiation of **13** and successive quench with *N,N'*-dimethylformamide (DMF) and triethylchlorosilane (TESCl) delivered the alkynyl aldehyde **14** in exquisite chemo- and regioselectivity (Fig. 3B). This intermediate was exposed to a vinyl zinc reagent derived from (*Z*)-1-bromo-2-ethoxyethene **15**, followed by elimination under acidic conditions with aqueous sodium bisulfate to give enal **17**. A catalytic enantioselective propar-

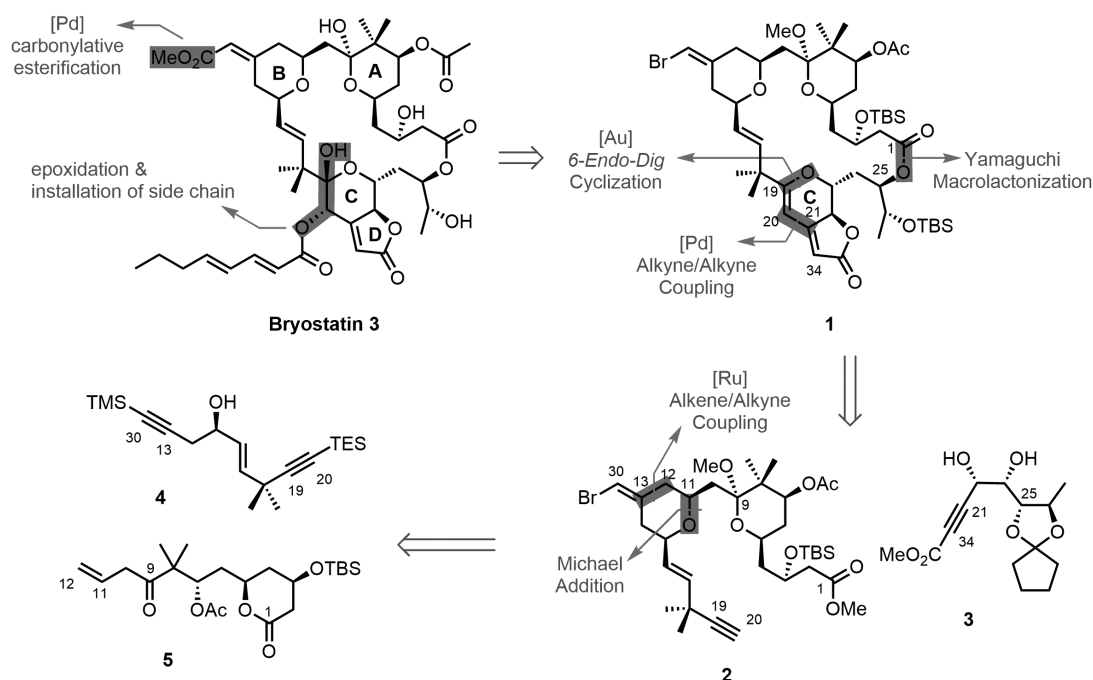
ylation of **17** afforded the diyne fragment **4** in 86% yield and 98% enantiomeric excess (36). The last key fragment **5** was prepared in eight steps according to Krische's procedure in their synthesis of Bryostatin 7 (20).

With fragments **3**, **4**, and **5** in hand, we proceeded with the fragment couplings and downstream elaborations. In the first key coupling, fragments **4** and **5** were assembled into tetrahydropyran **21** through a ruthenium-catalyzed alkene-alkyne coupling–Michael addition cascade (Fig. 3C) (37). The diyne fragment **4** underwent the desired coupling reaction with perfect chemoselectivity for the sterically more accessible alkyne. The cascade Michael addition reaction to construct the tetrahydropyran ring **B** was highly diastereoselective, with a >20/1 syn/anti ratio. The coupling product **21** was isolated in 48% yield (87% yield based on recovered starting material). Attempts to optimize the coupling reaction—including changing solvents [such as ethyl acetate, 3-pentanone, cyclopentanone, DMF, tetrahydrofuran (THF), and CH₂Cl₂], using additives [such as acetic acid, camphorsulfonic acid (CSA), acetonitrile, trimethylacetone, DMF, and 2,6-di-*tert*-butyl-4-methylpyridine], and varying reaction temperature and time—all led to detrimental effects. Fortunately, although the catalyst turnover was not ideal, the reaction was highly reproducible and easily performed at room temperature in air without precautions, and both unreacted coupling partners could be recovered in good yield and reused for multiple runs. The resulting vinyl silane **21** was *ipso*-brominated with complete retention of olefin geometry to furnish the vinyl bromide

22. Treatment of the latter with pyridinium *p*-toluenesulfonate (PPTS) in methanol afforded the bis-tetrahydropyran **23** through an acid-promoted ring-opening transesterification and ring-closing ketalization sequence. Such a cascade reaction was notably clean and chemoselective, with no transesterification of the C7 acetate. Intermediate **23** was then treated with silver(I) nitrate in aqueous THF to achieve a selective desilylation of the triethylsilyl (TES)-protected alkyne in the presence of the *tert*-butyldimethylsilyl (TBS)-protected C3 alcohol. The advanced intermediate **2** was isolated in 87% yield.

Intermediate **2** was then merged with fragment **3** by using a palladium-catalyzed alkyne-alkyne coupling reaction to give the transient intermediate **24** (Fig. 4) (38). The stereoselective nature of the coupling reaction accounted for the geometrically defined enoate of **24** which, in combination with the pendant-free diol, facilitated a spontaneous ring-closing transesterification to generate the requisite butenolide ring D of bryostatin 3 (39). The process was exclusively chemoselective, with no six-membered lactone formation, presumably owing to kinetic control of cyclization. The juxtaposition of the remaining free hydroxyl group and the alkyne deriving from the alkyne-alkyne coupling process in intermediate **25** set the stage for a 6-*endo-dig* cyclization to form the dihydropyran ring C in intermediate **26**. Thanks to the relatively acidic and mild conditions of the previous alkyne-alkyne coupling reaction, a stock solution of cationic gold catalyst could be directly injected to the reaction flask, without solvent exchange, to implement

Fig. 2. Retrosynthetic analysis of bryostatin 3.



the desired cyclization. After full conversion was indicated by means of thin-layer chromatography (TLC), a stock solution of ZrCl_4 in methanol was added to the above crude reaction mixture, again with no need of sol-

vent removal or exchange, to hydrolyze the C25-C26 cyclopentylidene acetal in situ. After the above three orthogonal operations in one pot, the triol **27** could be isolated in 48% overall yield. Although the C3-OTBS group was

desilylated during the acetone deprotection process, it was reinstalled by subjecting the triol to silylation conditions that resulted in bis-silylated product **28** in 60% yield. Moving forward with intermediate **28**, trimethyltin

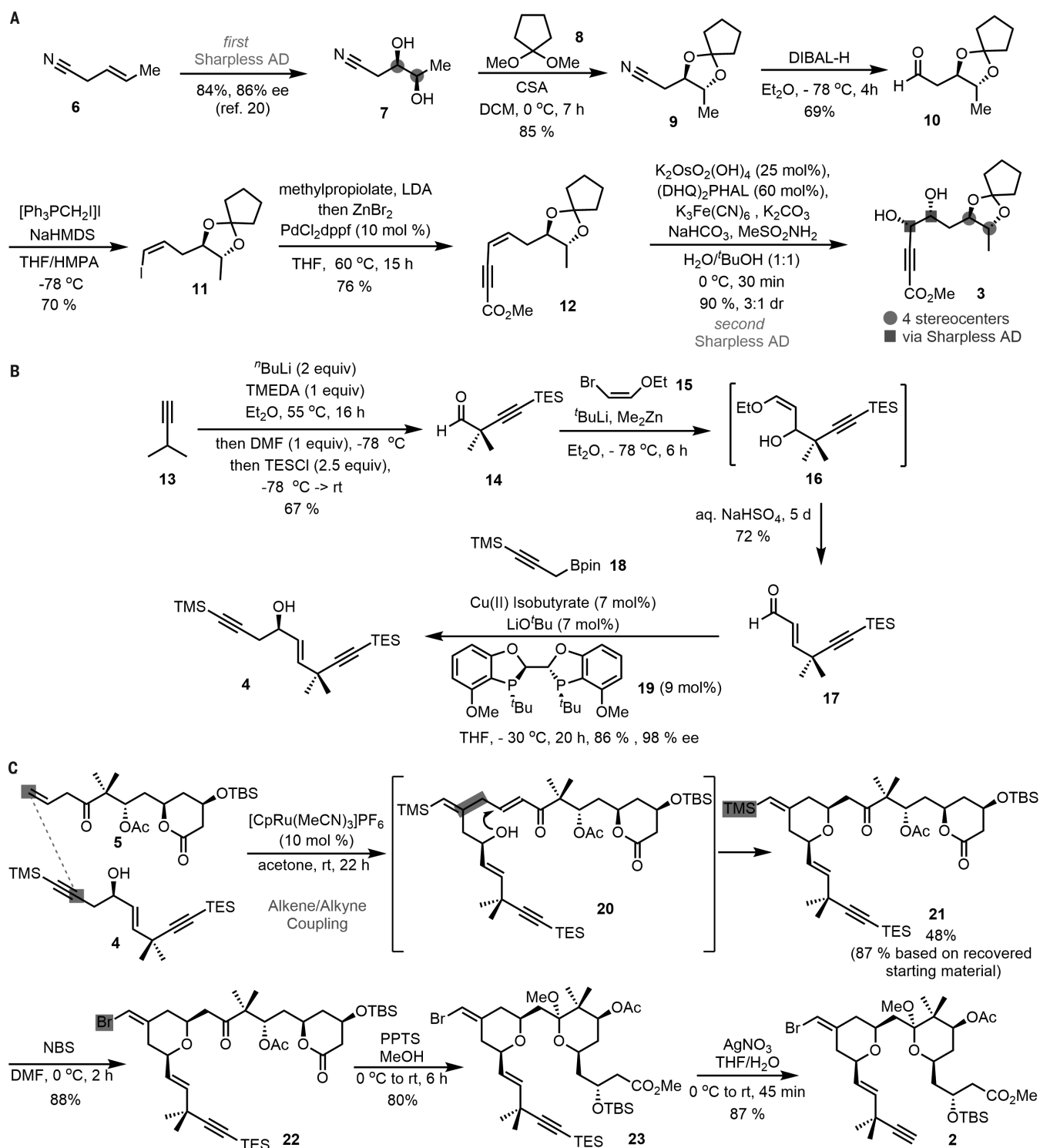


Fig. 3. Synthesis of **2, **3**, and **4**.** (A) Synthesis of fragment **3**. (B) Synthesis of fragment **4**. (C) Synthesis of intermediate **2**. CSA, camphorsulfonic acid; DIBAL-H, diisobutylaluminum hydride; NaHMDS, sodium hexamethyldisilazide; dppf, 1,1'-bis(diphenylphosphino)ferrocene; TMEDA, *N,N,N',N'*-tetramethylethylenediamine; DMF, *N,N*-dimethylformamide; TES, triethylsilyl; NBS, *N*-bromosuccinimide; PPTS, pyridinium *p*-toluenesulfonate; THF, tetrahydrofuran.

hydroxide selectively hydrolyzed the methyl ester to give *seco*-acid **29** in 80% yield (**40**), which was then cyclized under the modified Yamaguchi conditions to afford the macrocyclic intermediate **1** in near quantitative yield (**41**).

With an established route to the macrocyclic intermediate **1**, we proceeded to complete the total synthesis by means of an oxidative functionalization of the dihydropyran ring C, an esterification of the B-ring vinyl-bromide, and a global deprotection (Fig. 5). The dihydropyran ring C of intermediate **1** was chemoselectively epoxidized with methylrhenium trioxide/urea hydroperoxide (MTO/UHP) and 1-methylimidazole under the Yamazaki conditions (**42**). 1-Methylimidazole proved to be the superior additive for this process, presumably by buffering the Lewis acidity of MTO, accelerating the reaction and increasing the metal complex's lifetime. Although the reaction was relatively clean with high conversion according to nuclear magnetic resonance (NMR) monitoring of the crude product, the epoxide **30** was rather acid sensitive, and attempts to purify it only led to low and variable yields. Thus, after workup, the crude epoxide was directly subjected to the following acid-promoted ring-opening methanolysis step to furnish the α -hydroxyl ketal intermediate **31**, which was again unstable for purification and storage. Consequently, the crude α -hydroxyl ketal **31** was esterified immediately through acylation by using 2,4-octadienoic anhydride to give a stable intermediate **33**, which could be purified and stored regularly. The three-step, one-purification process with purification only at the end was highly reproducible and scalable. With the C-ring fully functionalized, we then turned to the carbonylation of the B-ring vinyl-bromide to install the requisite exocyclic vinyl ester. Treatment of the vinyl-bromide **33** with $\text{Pd}_2(\text{dba})_3 \cdot \text{CHCl}_3$ (20 mol %) and Xantphos (60 mol %) under one atmosphere of CO and with MeOH as a cosolvent furnished the desired vinyl ester **34** in 50% yield. At this stage, a protected form of bryostatin 3, **34**, was obtained. All that remained was deprotection by means of hydrolysis of the two methyl ketals and removal of the two silyl groups. Previously, Song's synthesis of bryostatin 8 (**21**) required the same type of final global deprotection. However, distinct from their synthesis in which aqueous HF in acetonitrile enabled the global deprotection in one pot, the treatment of our intermediate **34** under similar conditions only led to disappointingly complicated mixtures, as indicated by the messy NMR spectra and streaky TLC plates. Inspired by Yamamura's synthesis of bryostatin 3 (**18**), we performed the final global deprotection in two separate steps: The protected bryostatin **34** was first treated with aqueous hydrofluoride (HF) in acetonitrile to re-

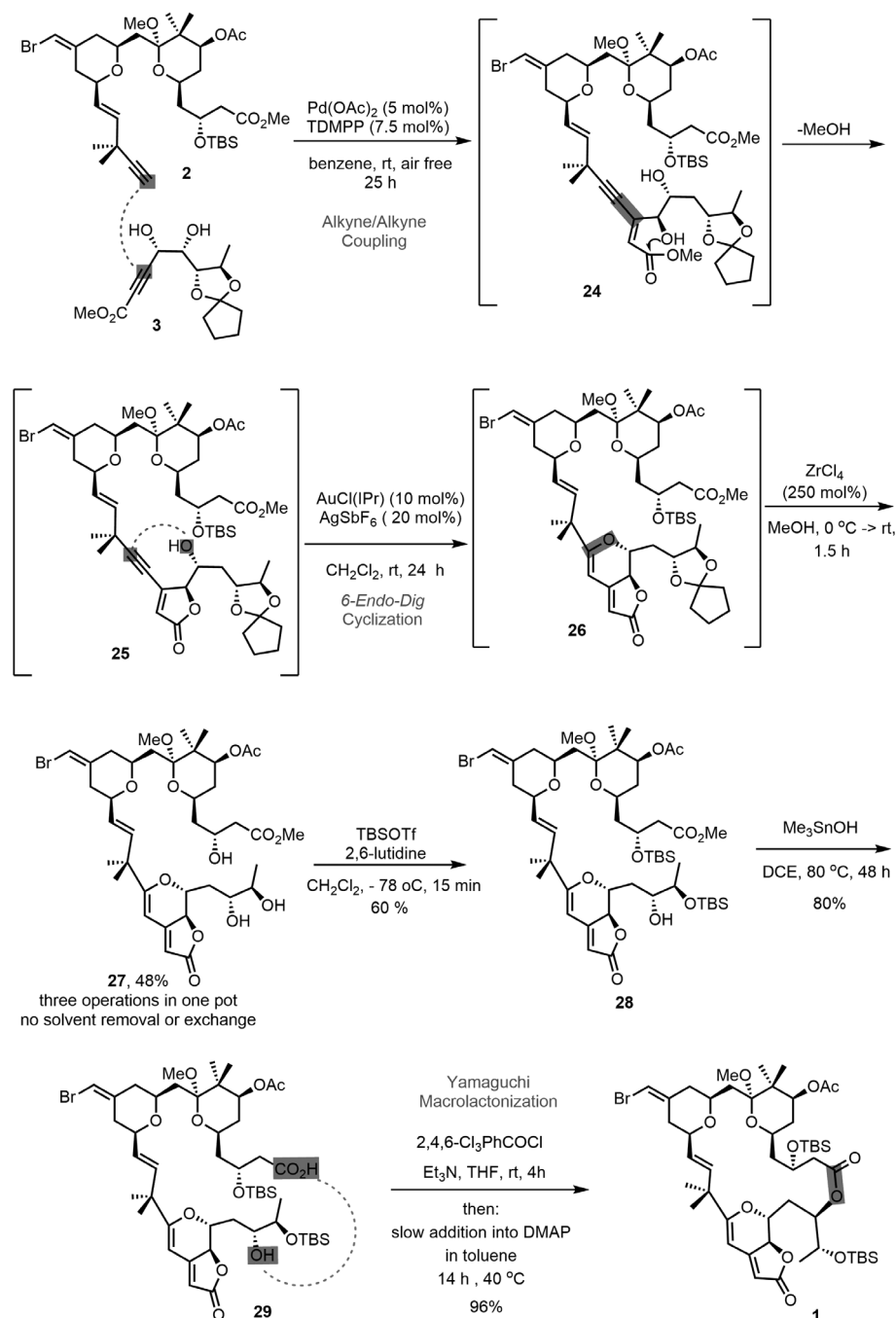


Fig. 4. Coupling of fragment 3 and intermediate 2 and further elaborations. TDMPP, tris(2,6-dimethoxyphenyl)phosphine; IPr, 1,3-bis(2,6-diisopropylphenyl)imidazol-2-ylidene; TBS, *tert*-butyldimethylsilyl; DCE, 1,2-dichloroethane; DMAP, 4-dimethylaminopyridine; THF, tetrahydrofuran.

move the two silyl groups and one methyl ketal. After workup, the crude residue was stirred in trifluoroacetic acid (TFA)- H_2O - CH_2Cl_2 for 1 hour to hydrolyze the remaining methyl ketal. Through the two-step deprotection procedure, 0.9 mg of bryostatin 3 was obtained in 60% overall yield from **34**.

The identification of our synthetic bryostatin 3 was initially problematic because

its ^1H -NMR was observed to be concentration dependent. Such a property has also been documented for other bryostatins (**15**, **16**, **43**). We were able to closely compare our synthetic bryostatin 3 with 1 mg of bryostatin 3 from K. Ohmori and found that the data of TLC, ^1H -NMR, ^{13}C -NMR, optical rotation, and high-performance liquid chromatography matched excellently (supplementary materials).

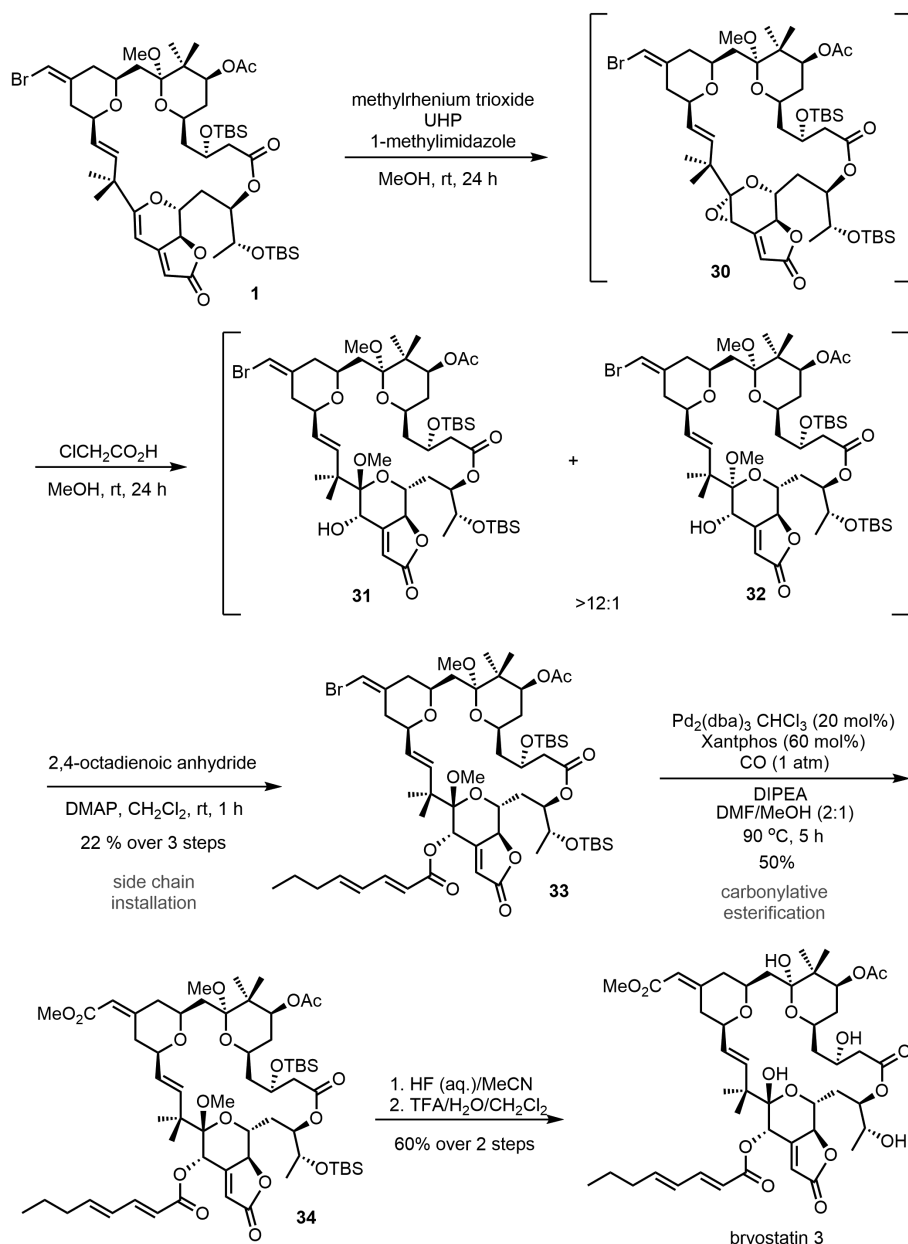


Fig. 5. Completion of broystatin 3 synthesis. UHP, urea hydroperoxide; DMAP, 4-dimethylaminopyridine; DIPEA, *N,N*-diisopropylethylamine; Xantphos, 9,9-dimethyl-4,5-bis(diphenylphosphino)xanthene; DMF, *N,N'*-dimethylformamide; TFA, trifluoroacetic acid.

This concise total synthesis showcases the value of the alkyne functionality that emanates from its chemoselectivity and its flexibility in elaborating the required structural details. The synthetic strategy and synthetic technologies in this work are expected to have implications toward the ultimate goal of practical, flexible, and concise synthesis of bryostatins and their bioactive analogs.

REFERENCES AND NOTES

1. G. R. Pettit, J. F. Day, J. L. Hartwell, H. B. Wood, *Nature* **227**, 962-963 (1970).
2. G. R. Pettit *et al.*, *J. Am. Chem. Soc.* **104**, 6846-6848 (1982).

- trioxide
azole
24 h
- 30
- 32
- >12:1
- 33
- Pd(dba)₃ CHCl₃ (20 mol%)
Xantphos (60 mol%)
CO (1 atm)
DIPEA
DMF/MeOH (2:1)
90 °C, 5 h
50%
carbonylative
esterification
- 34
- MeCN
O/CH₂Cl₂
2 steps
- bryostatin 3
13. B. P. Lucke-Wold et al., *Mol. Neurobiol.* **52**, 1219–1134 (2015).
14. B. A. Loy et al., *J. Am. Chem. Soc.* **137**, 3678–3685 (2015).
15. G. E. Keck, Y. B. Poudel, T. J. Cummins, A. Rudra, J. A. Covel, *J. Am. Chem. Soc.* **133**, 744–747 (2011).
16. P. A. Wender et al., *Science* **358**, 218–223 (2017).
17. D. A. Evans et al., *J. Am. Chem. Soc.* **121**, 7540–7552 (1999).
18. K. Ohmori et al., *Angew. Chem. Int. Ed.* **39**, 2290–2294 (2000).
19. M. Kageyama et al., *J. Am. Chem. Soc.* **112**, 7407–7408 (1990).
20. Y. Lu, S. K. Woo, M. J. Krische, *J. Am. Chem. Soc.* **133**, 13876–13879 (2011).
21. Y. Zhang et al., *Angew. Chem. Int. Ed.* **57**, 942–946 (2018).
22. P. A. Wender, A. J. Schrier, *J. Am. Chem. Soc.* **133**, 9228–9231 (2011).
23. B. M. Trost, G. Dong, *Nature* **456**, 485–488 (2008).
24. K. J. Hale, J. A. Lennon, S. Manaviazar, M. H. Javadi, C. J. Hobbs, *Tetrahedron Lett.* **36**, 1359–1362 (1995).
25. K. J. Hale, M. Frigerio, S. Manaviazar, *Org. Lett.* **5**, 503–505 (2003).
26. K. J. Hale, M. Frigerio, S. Manaviazar, *Org. Lett.* **3**, 3791–3794 (2001).
27. S. Manaviazar et al., *Org. Lett.* **8**, 4477–4480 (2006).
28. J. V. Allen et al., *Tetrahedron Lett.* **49**, 6352–6355 (2008).
29. B. J. De, M. Vandewalle, *Pure Appl. Chem.* **68**, 715–718 (2009).
30. R. Roy, A. W. Rey, M. Charron, R. Molino, *J. Chem. Soc. Chem. Commun.* **18**, 1308–1310 (1989).
31. E. A. Voight, H. Seradj, P. A. Roethlis, S. D. Burke, *Org. Lett.* **6**, 4045–4048 (2004).
32. J. S. Yadav, A. Bandyopadhyay, A. C. Kunwar, *Tetrahedron Lett.* **42**, 4907–4911 (2001).
33. P. A. Wender et al., *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7197–7201 (1988).
34. G. Stork, K. Zhao, *Tetrahedron Lett.* **30**, 2173–2174 (1989).
35. E. Negishi, M. Qian, F. Zeng, L. Anastasia, D. Babinski, *Org. Lett.* **5**, 1597–1600 (2003).
36. D. R. Fandrick et al., *J. Am. Chem. Soc.* **132**, 7600–7601 (2010).
37. B. M. Trost, H. Yang, G. Wuitschik, *Org. Lett.* **7**, 4761–4764 (2005).
38. B. M. Trost, M. T. Sorum, C. Chan, G. Rüther, *J. Am. Chem. Soc.* **119**, 698–708 (1997).
39. B. M. Trost, M. C. McIntosh, *J. Am. Chem. Soc.* **117**, 7255–7256 (1995).
40. K. C. Nicolaou, A. A. Estrada, M. Zak, S. H. Lee, B. S. Safina, *Angew. Chem. Int. Ed.* **44**, 1378–1382 (2005).
41. J. Inanaga, K. Hirata, H. Saeki, T. Katsuki, M. Yamaguchi, *Bull. Chem. Soc. Jpn.* **52**, 1989–1993 (1979).
42. S. Yamazaki, *Org. Biomol. Chem.* **8**, 2377–2385 (2010).
43. D. E. Schaufelberger, G. N. Chmurny, M. P. Koclek, *Magn. Reson. Chem.* **29**, 366–374 (1991).
- ACKNOWLEDGMENTS**
- S. Lynch is acknowledged for conducting two-dimensional NMR experiments. We thank K. Ohmori from the Tokyo Institute of Technology for the generous sharing of 1 mg of bryostatin 3.
- Funding:** We are grateful to the Tamaki Foundation for financial support. O.K. thanks Deutsche Forschungsgemeinschaft (DFG) for sponsoring the postdoctoral research. **Author contributions:** B.M.T. conceived and designed the project. Y.W., A.K.B., Z.H., M.H.N., and O.K. carried out the experiments. Y.W. and A.K.B. wrote the manuscript. B.M.T., Z.H., and J. S. Tracy proofread the manuscript. All authors analyzed the data and discussed the results. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.
- SUPPLEMENTARY MATERIALS**
- science.sciencemag.org/content/368/6494/1007/suppl/DC1
Materials and Methods
Supplementary Text
Tables S1 and S2
NMR Spectra
- 13 March 2020; accepted 22 April 2020
10.1126/science.abb7271

ACKNOWLEDGMENTS

S. Lynch is acknowledged for conducting two-dimensional NMR experiments. We thank K. Ohmori from the Tokyo Institute of Technology for the generous sharing of 1 mg of brivostatin 3.

Funding: We are grateful to the Tamaki Foundation for financial support. O.K. thanks Deutsche Forschungsgemeinschaft (DFG) for sponsoring the postdoctoral research. **Author contributions:** B.M.T. conceived and designed the project. Y.W., A.K.B., Z.H., M.H.N., and O.K. carried out the experiments. Y.W. and A.K.B. wrote the manuscript. B.M.T., Z.H., and J. S. Tracy proofread the manuscript. All authors analyzed the data and discussed the results. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text and the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6494/1007/suppl/DC1
Materials and Methods
Supplementary Text
Tables S1 and S2
NMR Spectra

13 March 2020; accepted 22 April 2020
10.1126/science.abb7271

CORONAVIRUS

Comparative pathogenesis of COVID-19, MERS, and SARS in a nonhuman primate model

Barry Rockx^{1*}, Thijs Kuiken¹, Sander Herfst¹, Theo Bestebroer¹, Mart M. Lamers¹, Bas B. Oude Munnink¹, Dennis de Meulder¹, Geert van Amerongen², Judith van den Brand^{1†}, Nisreen M. A. Okba¹, Debby Schipper¹, Peter van Run¹, Lonneke Leijten¹, Reina Sikkema¹, Ernst Verschoor³, Babs Verstrepen³, Willy Bogers³, Jan Langermans^{4,5}, Christian Drosten⁶, Martje Fentener van Vlissingen⁷, Ron Fouchier¹, Rik de Swart¹, Marion Koopmans¹, Bart L. Haagmans^{1*}

The current pandemic coronavirus, severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2), was recently identified in patients with an acute respiratory syndrome, coronavirus disease 2019 (COVID-19). To compare its pathogenesis with that of previously emerging coronaviruses, we inoculated cynomolgus macaques with SARS-CoV-2 or Middle East respiratory syndrome (MERS)–CoV and compared the pathology and virology with historical reports of SARS-CoV infections. In SARS-CoV-2–infected macaques, virus was excreted from nose and throat in the absence of clinical signs and detected in type I and II pneumocytes in foci of diffuse alveolar damage and in ciliated epithelial cells of nasal, bronchial, and bronchiolar mucosae. In SARS-CoV infection, lung lesions were typically more severe, whereas they were milder in MERS-CoV infection, where virus was detected mainly in type II pneumocytes. These data show that SARS-CoV-2 causes COVID-19–like disease in macaques and provides a new model to test preventive and therapeutic strategies.

After the first reports of an outbreak of an acute respiratory syndrome in China in December 2019, a novel coronavirus, severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2), was identified (1, 2). As of 14 March 2020, over 140,000 cases were reported worldwide with more than 5400 deaths, surpassing the combined number of cases and deaths of two previously emerging coronaviruses, SARS-CoV and Middle East respiratory syndrome (MERS)–CoV (3). The disease caused by this virus, coronavirus disease 2019 (COVID-19), is characterized by a range of symptoms, including fever, cough, dyspnoea, and myalgia in most cases (2). In severe cases, bilateral lung involvement with ground-glass opacity is the most common chest computed tomography (CT) finding (4). Similarly to the 2002–2003 outbreak of SARS, the severity of COVID-19 disease is associated with increased age and/or a comorbidity, although severe disease is not limited to these risk groups (5). However, despite the large number of cases and deaths, limited information is available on the pathogenesis of this virus infection.

Two reports on the histological examination of the lungs of three patients showed bilateral diffuse alveolar damage (DAD), pulmonary edema, and hyaline membrane formation, indicative of acute respiratory distress syndrome (ARDS), as well as characteristic syncytial cells in the alveolar lumen (6, 7), similar to findings during the 2002–2003 outbreak of SARS-CoV (8). The pathogenesis of SARS-CoV infection was previously studied in a nonhuman primate model (cynomolgus macaques) where aged animals were more likely to develop disease (9–13). In the current study, SARS-CoV-2 infection was characterized in the same animal model and compared with infection with MERS-CoV and historical data on SARS-CoV (9, 10, 12).

First, two groups of four cynomolgus macaques [both young adult (young), 4 to 5 years of age; and old adult (aged), 15 to 20 years of age] were inoculated by a combined intratracheal (IT) and intranasal (IN) route with a SARS-CoV-2 strain from a German traveler returning from China. No overt clinical signs were observed in any of the infected animals, except for a serous nasal discharge in one aged animal on day 14 post inoculation (p.i.). No significant weight loss was observed in any of the animals during the study. By day 14 p.i., all remaining animals seroconverted as revealed by the presence of SARS-CoV-2–specific antibodies against the virus S1 domain and nucleocapsid proteins in their sera (fig. S1).

As a measure of virus shedding, nasal, throat, and rectal swabs were assayed for virus by reverse transcription–quantitative polymerase chain reaction (RT–qPCR) and virus culture. In nasal swabs, SARS-CoV-2 RNA peaked by day 2 p.i. in young animals and by day 4 p.i. in

aged animals and was detected up to at least day 8 p.i. in two out of four animals and up to day 21 p.i. in one out of four animals (Fig. 1A). Overall, higher levels of SARS-CoV-2 RNA were detected in nasal swabs of aged animals compared with young animals. SARS-CoV-2 RNA in throat swabs peaked at day 1 p.i. in young and day 4 p.i. in aged animals and decreased more rapidly over time by comparison with the nasal swabs but could still be detected intermittently up to day 10 p.i. (Fig. 1B). Low levels of infectious virus were cultured from throat and nasal swabs up to day 2 and 4, respectively (table S1). In support of virus shedding by these animals, environmental sampling was performed to determine potential contamination of surfaces. Environmental sampling indicated the presence of low levels of SARS-CoV-2 RNA on surfaces through both direct contact (hands) and indirect contamination within the isolator (table S2). SARS-CoV-2 RNA was only detected in a rectal swab from one animal on day 14 p.i., and no viral RNA was detected in whole blood at any time point throughout the study.

On autopsy of four macaques on day 4 p.i., two had foci of pulmonary consolidation in the lungs (Fig. 2A). One animal (aged: 17 years) showed consolidation in the right middle lobe, representing less than 5% of the lung tissue. A second animal (young: 5 years) had two foci in the left lower lobe, representing about 10% of the lung tissue (Fig. 2A). The consolidated lung tissue was well circumscribed, red-purple, level, and less buoyant than normal. The other organs in these two macaques, as well as the respiratory tract and other organs of the other two animals, were normal.

Virus replication was assessed by RT–qPCR on day 4 p.i. in tissues from the respiratory, digestive, urinary, and cardiovascular tracts and from endocrine and central nervous systems, as well as from various lymphoid tissues. Virus replication was primarily restricted to the respiratory tract (nasal cavity, trachea, bronchi, and lung lobes) with highest levels of SARS-CoV-2 RNA in lungs (Fig. 1C). In three out of four animals, SARS-CoV-2 RNA was also detected in ileum and tracheo-bronchial lymph nodes (Fig. 1C).

The main histological lesion in the consolidated pulmonary tissues of both the young and aged animals involved the alveoli and bronchioles and consisted of areas with acute or more advanced DAD (Fig. 2B). In these areas, the lumina of alveoli and bronchioles were variably filled with protein-rich edema fluid, fibrin, and cellular debris; alveolar macrophages; and fewer neutrophils and lymphocytes (Fig. 2, C to E). There was epithelial necrosis with extensive loss of epithelium from alveolar and bronchiolar walls. Hyaline membranes were present in a few damaged alveoli. In areas with more advanced lesions, the alveolar walls

¹Department of Viroscience, Erasmus University Medical Center, Rotterdam, Netherlands. ²Viroclinics Xplore, Schaijk, Netherlands. ³Department of Virology, Biomedical Primate Research Centre, Rijswijk, Netherlands. ⁴Animal Science Department, Biomedical Primate Research Centre, Rijswijk, Netherlands. ⁵Population Health Sciences, Unit Animals in Science and Society, Faculty of Veterinary Medicine, Utrecht University, Netherlands. ⁶Institute of Virology, Charité–Universitätsmedizin, Berlin, Germany. ⁷Erasmus Laboratory Animal Science Center, Erasmus University Medical Center, Rotterdam, Netherlands.

*Corresponding author. Email: b.rockx@erasmusmc.nl (B.R.); b.haagmans@erasmusmc.nl (B.L.H.) †Present address: Division of Pathology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, Netherlands

were moderately thickened and lined by cuboidal epithelial cells (type II pneumocyte hyperplasia), and the alveolar lumina were empty. Alveolar and bronchiolar walls were thickened by edema fluid, mononuclear cells, and neu-

trophils. There were aggregates of lymphocytes around small pulmonary vessels. Moderate numbers of lymphocytes and macrophages were present in the lamina propria and submucosa of the bronchial walls, and a few neutrophils

were detected in the bronchial epithelium. Regeneration of epithelium was seen in some bronchioles, visible as an irregular layer of squamous to high cuboidal epithelial cells with hyperchromatic nuclei. There were occasional

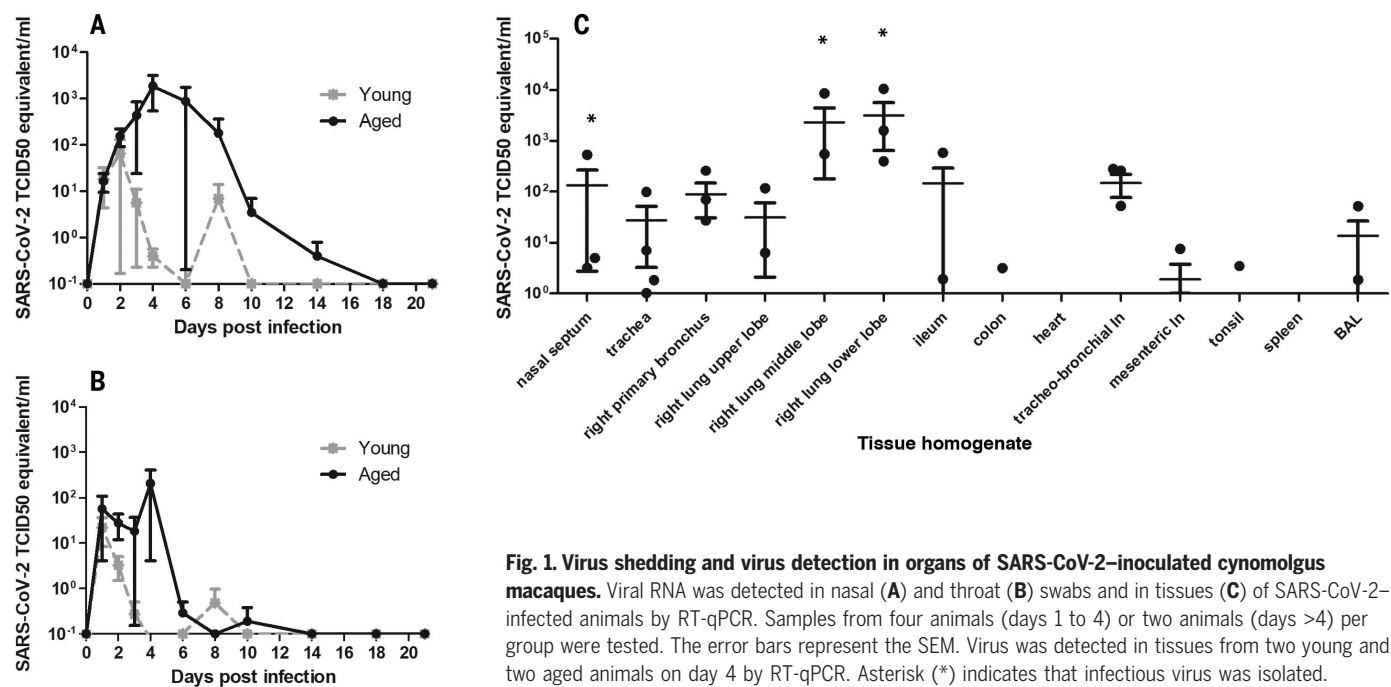


Fig. 1. Virus shedding and virus detection in organs of SARS-CoV-2-inoculated cynomolgus macaques. Viral RNA was detected in nasal (A) and throat (B) swabs and in tissues (C) of SARS-CoV-2-infected animals by RT-qPCR. Samples from four animals (days 1 to 4) or two animals (days >4) per group were tested. The error bars represent the SEM. Virus was detected in tissues from two young and two aged animals on day 4 by RT-qPCR. Asterisk (*) indicates that infectious virus was isolated.

Table 1. Comparative pathogenesis of SARS-CoV-2, MERS-CoV, and SARS-CoV infections in cynomolgus macaques. Max, maximum; Ref., reference.													
Virus	Age category	No.	Clinical signs	Max. excretion from throat*	Max. excretion from nose*	Viral titer in lung†	Virus in extra-respiratory tissues	Pulmonary lesions at 4 days post inoculation					Ref.
								% affected lung	Hyaline membranes	Alveolar edema	Leuko-cyte infiltration	Cell type tropism	
SARS-CoV	Young	4	No	10e3.5	10e4.2	10e6.5	No	0–5	No	No	Yes	Type I & II pneumocytes	(10–12)
SARS-CoV	Aged	4	Yes	10e3.0	10e3.0	10e6.2	Kidney	0–60	Yes	Yes	Yes	Type I & II pneumocytes	(10–12)
MERS-CoV	Young	4	No	10e5.3	10e4.3	10e4.4	Spleen	0–5	No	Yes (small amount)	Yes	Type II pneumocytes	This study
SARS-CoV-2	Young	2	No	10e1.8	10e2.4	10e4.0	Ileum, colon, tonsil	0–10	Yes	Yes	Yes	Type I & II pneumocytes	This study
SARS-CoV-2	Aged	2	No	10e2.9	10e3.7	10e3.1	Ileum, colon, tonsil	0–5	No	No	Yes	Type I & II pneumocytes	This study

*TCID₅₀eq per ml. †TCID₅₀eq per gram of tissue.

multinucleated giant cells (syncytia) free in the lumina of bronchioles and alveoli (Fig. 2F) and, based on positive pan-keratin staining and negative CD68 staining, these appeared to originate from epithelial cells (Fig. 2F, inset).

SARS-CoV-2 antigen expression was detected in moderate numbers of type I pneumocytes and a few type II pneumocytes within foci of DAD (Fig. 2, G and H, and fig. S2). The pattern of staining was similar to that in lung tissue from SARS-CoV-infected macaques (positive control). SARS-CoV-2 antigen expression was not observed in any of the syncytia. In addition, SARS-CoV-2 antigen expression

was detected in nonlesional tissues of all lung lobes in three out of four macaques (both young and one aged) in a few type I and II pneumocytes, bronchial ciliated epithelial cells, and bronchiolar ciliated epithelial cells. The other aged macaque, without virological or pathological evidence of SARS-CoV-2 infection in the lungs, did have SARS-CoV-2 antigen expression in ciliated epithelial cells of nasal septum (Fig. 2I), nasal concha, and palatum molle, in the absence of associated histopathological changes. No SARS-CoV antigen expression was detected in other sampled tissues, including brain and intestine.

To assess the severity of infection with SARS-CoV-2 compared with MERS-CoV, we inoculated young cynomolgus macaques (3 to 5 years of age) with MERS-CoV via the IN and IT route. All animals remained free of clinical signs. At day 21 p.i., all remaining animals ($n = 2$) seroconverted as revealed by the presence of MERS-CoV-specific antibodies in their sera by ELISA (fig. S3).

MERS-CoV RNA was detected in nasal (Fig. 3A) and throat swabs (Fig. 3B) on days 1 to 11 p.i., with peaks on days 1 and 2 p.i., respectively. Low levels [between 1 and 85 median tissue culture infectious dose (TCID₅₀) equivalent/ml]

Fig. 2. Characteristic pathological changes and virus antigen expression in the lungs of SARS-CoV-2-inoculated cynomolgus macaques.

(A) Two foci of pulmonary consolidation in the left lower lung lobe (arrowheads). (B) Area of pneumonia [staining with hematoxylin and eosin (H&E); bar, 0.5 cm]. (C) Edema fluid in alveolar lumina (H&E; bar, 25 μ m). (D) Neutrophils, as well as erythrocytes, fibrin, and cell debris, in an alveolar lumen flooded by edema fluid (H&E; bar, 10 μ m). (E) Mononuclear cells, either type II pneumocytes or alveolar macrophages, in an alveolar lumen flooded by edema fluid (H&E; bar, 10 μ m). (F) Syncytium in an alveolar lumen (H&E; 100 $\times$ objective). Inset: Syncytium expresses keratin, indicating epithelial cell origin [immunohistochemistry (IHC) for pankeratin AE1/AE3; bar, 10 μ m]. (G) SARS-CoV-2 antigen expression is colocalized with areas of diffuse alveolar damage (IHC for SARS-CoV-nucleocapsid; bar, 50 μ m). (H) Type I (flat) and type II (cuboidal) pneumocytes in affected lung tissue express SARS-CoV-2 antigen (IHC for SARS-CoV-nucleocapsid; bar, 25 μ m). (I) Ciliated columnar epithelial cells of respiratory mucosa in nasal cavity express SARS-CoV-2 antigen (IHC for SARS-CoV-nucleocapsid; bar, 25 μ m).

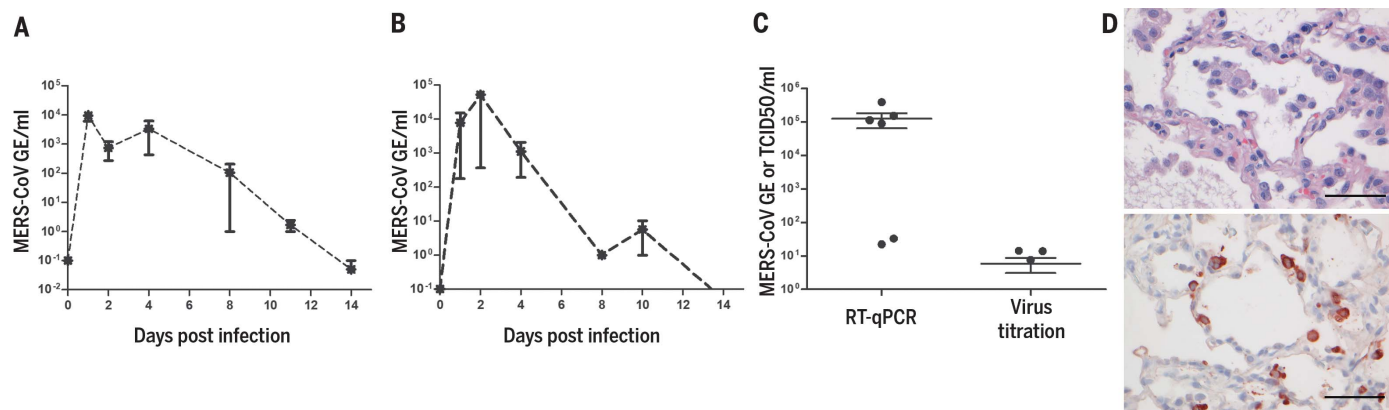
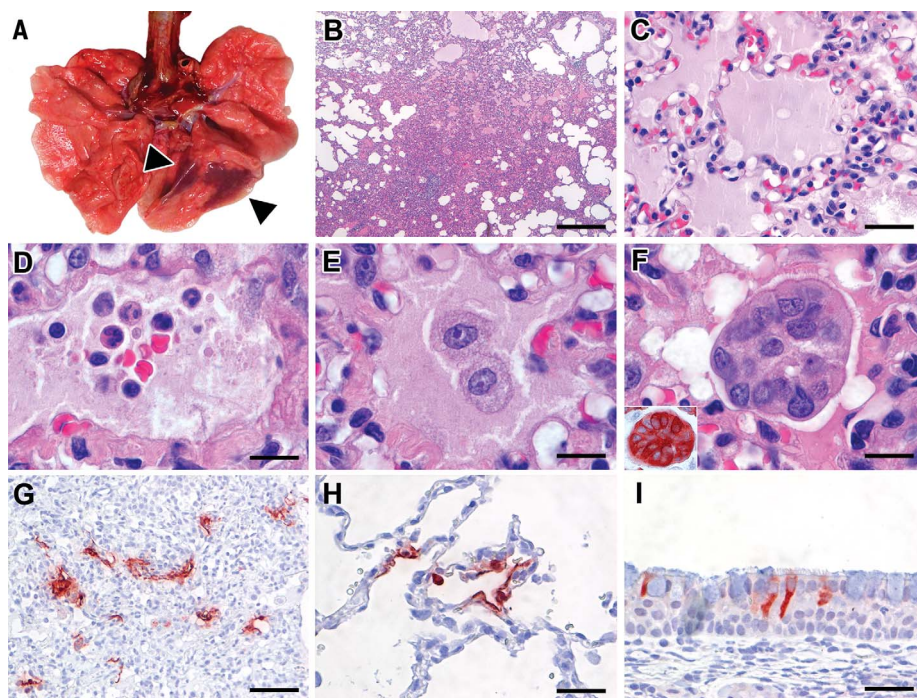


Fig. 3. Virus shedding and virus detection in organs of MERS-CoV-inoculated cynomolgus macaques. Viral RNA was detected in nasal (A) and throat (B) swabs and tissues (C) of MERS-CoV-infected animals by RT-qPCR. Samples from four animals per group were tested. The error bars represent the SEM. Virus was detected in tissues on day 4 by RT-qPCR. Histopathological changes (D) (left) with hypertrophic and hyperplastic type II pneumocytes in the alveolar septa and increased numbers of alveolar macrophages in the alveolar lumina and virus antigen expression (right) in type II pneumocytes. Bar, 50 μ m.

of MERS-CoV RNA were detected in rectal swabs on days 2 and 3 p.i.

At autopsy of four macaques at day 4 p.i., three animals had foci of pulmonary consolidation, characterized by slightly depressed areas in the lungs, representing less than 5% of the lung tissue (Table 1). Similar to SARS-CoV-2 infection in both young and aged animals, on day 4 p.i., MERS-CoV RNA was primarily detected in the respiratory tract of inoculated animals (Fig. 3C). Infectious virus titers were comparable to those of SARS-CoV-2 infection, but lower compared to SARS-CoV infection, of young macaques (Table 1). In addition, MERS-CoV RNA was detected in the spleen (Table 1).

Consistent with the presence of virus in the lower respiratory tract at day 4 p.i., histopathological changes characteristic for DAD were observed in the lungs of inoculated animals (Fig. 3D). The alveolar septa were thickened owing to infiltration of neutrophils and macrophages, and to moderate type II pneumocyte hyperplasia and hypertrophy. In the alveolar lumina, there were increased numbers of alveolar macrophages and some edema fluid containing fibrin and some neutrophils (Fig. 3D). Few syncytial cells were seen in the alveolar lumina. MERS-CoV antigen was not detected in tissues on day 4 p.i. in any part of the respiratory tract. We therefore sampled four young macaques at day 1 p.i. At this time, we observed multifocal expression of viral antigen, predominantly in type II pneumocytes and occasionally in type I pneumocytes, bronchiolar and bronchial epithelial cells, and some macrophages (Fig. 3D).

In summary, we inoculated young and aged cynomolgus macaques with a low-passage clinical isolate of SARS-CoV-2, which resulted in productive infection in the absence of overt clinical signs. Recent studies in human cases have shown that presymptomatic and asymptomatic cases can also shed virus (14, 15). Increased age did not affect disease outcome, but there was prolonged viral shedding in the upper respiratory tract of aged animals. Prolonged shedding has been observed in both SARS-CoV-2 and SARS-CoV patients (16, 17). SARS-CoV-2 shedding in our asymptomatic model peaked early in the course of infection, similar to what is seen in symptomatic patients (16). Also, SARS-CoV-2 antigen was detected in ciliated epithelial cells of nasal mucosae at day 4 p.i., which was not seen for SARS-CoV (10) or MERS-CoV infections (this study) in this animal model. Viral tropism for the nasal mucosa fits with efficient respiratory transmission, as has been seen for influenza A virus (18). This early peak in virus shedding for SARS-CoV-2 is similar to that of influenza virus shedding (19) and may explain why case detection and isolation may not be as effective for SARS-CoV-2 as it was for the control of SARS-CoV (20). SARS-CoV-2 was

primarily detected in tissues of the respiratory tract; however, SARS-CoV-2 RNA was also detectable in other tissues such as intestines, in line with a recent report (21). Similar results regarding viral shedding and tissue and cell tropism were recently also reported after SARS-CoV-2 inoculation in rhesus macaques. However, unlike in our model, SARS-CoV-2 infection in rhesus macaques does result in transient respiratory disease and weight loss (22, 23).

Two out of four animals had foci of DAD on day 4 p.i. The colocalization of SARS-CoV-2 antigen expression and DAD provides strong evidence that SARS-CoV-2 infection caused this lesion. The histological character of the DAD, including alveolar and bronchiolar epithelial necrosis, alveolar edema, hyaline membrane formation, and accumulation of neutrophils, macrophages, and lymphocytes, corresponds with the limited pathological analyses of human COVID-19 cases (6, 7). In particular, the presence of syncytia in the lung lesions is characteristic of respiratory coronavirus infections. Whereas MERS-CoV primarily infects type II pneumocytes in cynomolgus macaques, both SARS-CoV and SARS-CoV-2 also infect type I pneumocytes. Injury to type I pneumocytes can result in pulmonary edema, and formation of hyaline membranes (24), which may explain why hyaline membrane formation is a hallmark of SARS and COVID-19 (7, 10) but not frequently reported for MERS (25, 26).

These data show that cynomolgus macaques are permissive to SARS-CoV-2 infection, shed virus for a prolonged period of time, and display COVID-19-like disease. In this non-human primate model, SARS-CoV-2 replicates efficiently in respiratory epithelial cells throughout the respiratory tract, including nasal cavity, bronchi, bronchioles, and alveoli. Replication in the upper respiratory tract fits with efficient transmission between hosts, whereas replication in the lower respiratory tract fits with the development of lung disease. An in-depth comparison of infection with SARS-CoV, MERS-CoV, and SARS-CoV-2 in this model may identify key pathways in the pathogenesis of these emerging viruses. This study provides a new infection model that will be critical in the evaluation and licensure of preventive and therapeutic strategies against SARS-CoV-2 infection for use in humans, as well as for evaluating the efficacy of repurposing species-specific existing treatments, such as pegylated interferon (12).

REFERENCES AND NOTES

1. J. F. Chan et al., *Lancet* **395**, 514–523 (2020).
2. C. Huang et al., *Lancet* **395**, 497–506 (2020).
3. E. Mahase, *BMJ* **368**, m641 (2020).
4. H. Shi et al., *Lancet Infect. Dis.* **20**, 425–434 (2020).
5. N. Chen et al., *Lancet* **395**, 507–513 (2020).
6. Z. Xu et al., *Lancet Respir. Med.* **8**, 420–422 (2020).

7. S. Tian et al., *J. Thorac. Oncol.* **S1556-0864(20)**30132-5 (2020).
8. J. M. Nicholls et al., *Lancet* **361**, 1773–1778 (2003).
9. R. A. Fouchier et al., *Nature* **423**, 240 (2003).
10. T. Kuiken et al., *Lancet* **362**, 263–270 (2003).
11. S. L. Smits et al., *PLOS Pathog.* **6**, e1000756 (2010).
12. B. L. Haagmans et al., *Nat. Med.* **10**, 290–293 (2004).
13. B. Rockx et al., *PLOS ONE* **6**, e18558 (2011).
14. C. Li et al., *Emerg. Infect. Dis.* **26** (2020).
15. A. Kimball et al., *MMWR Morb. Mortal. Wkly. Rep.* **69**, 377–381 (2020).
16. L. Zou et al., *N. Engl. J. Med.* **382**, 1177–1179 (2020).
17. J. S. Peiris et al., *Lancet* **361**, 1767–1772 (2003).
18. M. Richard et al., *Nat. Commun.* **11**, 766 (2020).
19. C. C. Li et al., *Emerg. Infect. Dis.* **16**, 1265–1272 (2010).
20. X. Pang et al., *JAMA* **290**, 3215–3221 (2003).
21. F. Xiao et al., *Gastroenterology* **S0016-5085(20)**30282-1 (2020).
22. W. D. Linlin Bao et al., Reinfection could not occur in SARS-CoV-2 infected rhesus macaques. *bioRxiv* 2020.03.13.990226 (2020). <https://doi.org/10.1101/2020.03.13.990226>.
23. V. J. Munster et al., Respiratory disease and virus shedding in rhesus macaques inoculated with SARS-CoV-2. *bioRxiv* 2020.03.21.001628 (2020). <https://doi.org/10.1101/2020.03.21.001628>.
24. L. B. Ware, M. A. Matthay, *N. Engl. J. Med.* **342**, 1334–1349 (2000).
25. W. J. Shieh et al., *Hum. Pathol.* **36**, 303–309 (2005).
26. D. L. Ng et al., *Am. J. Pathol.* **186**, 652–658 (2016).

ACKNOWLEDGMENTS

We thank L. de Meulder, A. van der Linden, I. Chestakova, and F. van der Panne for technical assistance. We also thank Y. Kap, D. Akkermans, V. Vaes, M. Sommers, and F. Meijers for assistance with the animal studies. **Funding:** This research is (partly) financed by the NWO Stevin Prize awarded to M.K. by the Netherlands Organisation for Scientific Research (NWO), H2020 grant agreement 874735-VEO to M.K., NIH/NIAD contract HHSN272201400008C to R.F., and H2020 grant agreement 101003651-MANCO to B.H. **Author contributions:** Conceptualization, B.R., T.K., R.F., M.K., R.d.S., M.F. B.H.; investigation, B.R., T.K., S.H., T.B., M.L., D.d.M., G.v.A., J.v.d.B., N.O., D.S., P.v.R., L.L.; resources, B.H., C.D., E.V., B.V., J.L.; supervision, B.R. and B.H.; writing, original draft, B.R., T.K., and B.H.; writing—review and editing, all authors; funding acquisition: S.H., B.H., R.F., and M.K. **Animal welfare:** Research was conducted in compliance with the Dutch legislation for the protection of animals used for scientific purposes (2014, implementing EU Directive 2010/63) and other relevant regulations. The licensed establishment where this research was conducted (Erasmus MC) has an approved OLAW Assurance no. A5051-01. Research was conducted under a project license from the Dutch competent authority and the study protocol (no. 17-4312) was approved by the institutional Animal Welfare Body. **Competing interests:** T.B., R.F., and B.H. are listed as inventors on a patent application related to MERS-coronavirus diagnostics, therapy and vaccines. Other authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6494/1012/suppl/DC1
Materials and Methods
Figs. S1 to S3
Tables S1 and S2
References (27–30)
MDAR Reproducibility Checklist

15 March 2020; accepted 15 April 2020
Published online 17 April 2020
10.1126/science.abb7314

CORONAVIRUS

Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS–coronavirus 2

Jianzhong Shi^{1*}, Zhiyuan Wen^{1*}, Gongxun Zhong^{1*}, Huanliang Yang^{1*}, Chong Wang^{1*}, Baoying Huang^{2*}, Renqiang Liu¹, Xijun He³, Lei Shuai¹, Ziruo Sun¹, Yubo Zhao¹, Peipei Liu², Libin Liang¹, Pengfei Cui¹, Jinliang Wang¹, Xianfeng Zhang³, Yuntao Guan³, Wenjie Tan², Guizhen Wu^{2†}, Hualan Chen^{1†}, Zhigao Bu^{1,3†}

Severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2) causes the infectious disease COVID-19 (coronavirus disease 2019), which was first reported in Wuhan, China, in December 2019. Despite extensive efforts to control the disease, COVID-19 has now spread to more than 100 countries and caused a global pandemic. SARS-CoV-2 is thought to have originated in bats; however, the intermediate animal sources of the virus are unknown. In this study, we investigated the susceptibility of ferrets and animals in close contact with humans to SARS-CoV-2. We found that SARS-CoV-2 replicates poorly in dogs, pigs, chickens, and ducks, but ferrets and cats are permissive to infection. Additionally, cats are susceptible to airborne transmission. Our study provides insights into the animal models for SARS-CoV-2 and animal management for COVID-19 control.

In late December 2019, an unusual pneumonia emerged in humans in Wuhan, China, and rapidly spread internationally, raising global public health concerns. The causative pathogen was identified as a novel coronavirus (1–16) and named severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2) on the basis of a phylogenetic analysis of related

coronaviruses by the Coronaviridae Study Group of the International Committee on Taxonomy of Viruses (17). Subsequently, the disease caused by this virus was designated coronavirus disease 2019 (COVID-19) by the World Health Organization (WHO). Despite major efforts to control the COVID-19 outbreak, the disease is still spreading. As of 11 March

2020, SARS-CoV-2 infections have been reported in more than 100 countries, and 118,326 human cases have been confirmed, with 4292 fatalities (18). WHO has now officially declared COVID-19 a pandemic.

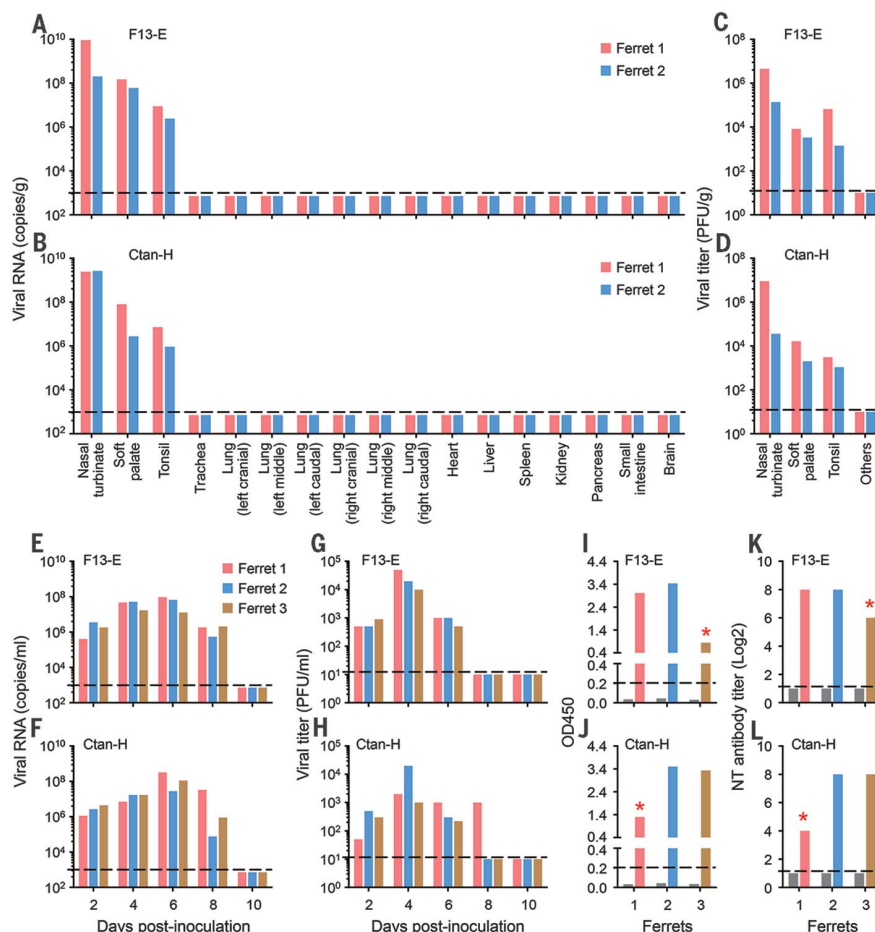
Although SARS-CoV-2 shares 96.2% of its identity at the nucleotide level with the coronavirus RaTG13—which was detected in horseshoe bats (*Rhinolophus* spp.) in Yunnan province, China, in 2013 (3)—it has not previously been detected in humans or other animals. The emerging public health crisis raises many urgent questions. Could the widely disseminated SARS-CoV-2 be transmitted to other animal species, which then become reservoirs of infection? The SARS-CoV-2 infection has a wide clinical spectrum in humans, ranging from mild infection to death, but how does

¹State Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Harbin 150069, People's Republic of China. ²National Institute for Viral Disease Control and Prevention, China CDC, Beijing 102206, People's Republic of China. ³National High Containment Laboratory for Animal Diseases Control and Prevention, Harbin 150069, People's Republic of China.

*These authors contributed equally to this work.

†Corresponding author. Email: buzhihao@caas.cn (Z.B.); chenhuailan@caas.cn (H.C.); wugz@ivdc.chinacdc.cn (G.W.)

Fig. 1. Replication of SARS-CoV-2 in ferrets. Viral RNA in organs or tissues of ferrets inoculated with (A) F13-E virus strain or (B) CTan-H strain. Viral titers in organs or tissues of ferrets inoculated with F13-E (C) and CTan-H (D). The viral RNA–negative organs indicated in (A) and (B) were also negative for virus titration [indicated as “Others” in (C) and (D)]. Viral RNA (E and F) and viral titer (G and H) in nasal washes of ferrets inoculated with F13-E [(E) and (G)] and CTan-H [(F) and (H)]. Antibodies against SARS-CoV-2 tested by ELISA (I and J) and neutralization assay (K and L) with sera derived from ferrets inoculated with F13-E [(I) and (K)] and CTan-H [(J) and (L)]. Each color bar represents the value from an individual animal. The gray bars in (I) to (L) indicate the antibody values of sera collected from each animal before inoculation. Asterisks denote animals that were euthanized on day 13 p.i.; the other four animals were euthanized on day 20 p.i. The dashed lines in (I) and (L) show the cutoff value for seroconversion, and the dashed lines in the other panels indicate the lower limit of detection. OD450, optical density measured at 450 nm; NT antibody, neutralizing antibody.



the virus behave in other animals? As efforts progress toward vaccine and antiviral drug development, which animal(s) can be used to most accurately model the efficacy of such control measures in humans? To address these questions, we evaluated the susceptibility of different model laboratory animals, as well as companion and domestic animals, to SARS-CoV-2.

All experiments with infectious SARS-CoV-2 were performed in the biosafety level 4 and animal biosafety level 4 facilities in the Harbin Veterinary Research Institute (HVRI) of the Chinese Academy of Agricultural Sciences (CAAS), which was approved for such use by the Ministry of Agriculture and Rural Affairs of China. Details of the biosafety and biosecurity measures are provided in the supplementary materials (19). The protocols for animal study and animal welfare were reviewed and approved by the Committee on the Ethics of Animal Experiments of the HVRI of CAAS (approval number 2020-01-01JiPi).

Ferrets are commonly used as an animal model for viral respiratory infections in humans (20–26). We therefore tested the susceptibility of ferrets to SARS-CoV-2. Two virus strains were used in this study: (i) SARS-CoV-2/F13/environment/2020/Wuhan (F13-E), isolated from an environmental sample collected in the Huanan Seafood Market in Wuhan, and (ii) SARS-CoV-2/CTan/human/2020/Wuhan (CTan-H), isolated from a human patient. Pairs of ferrets were inoculated intranasally with 10^5 plaque-forming units (PFU) of F13-E or CTan-H and euthanized on day 4 post-inoculation (p.i.). The nasal turbinate, soft palate, tonsils, trachea, lung, heart, spleen, kidneys, pancreas, small intestine, and brain from each ferret were collected for viral RNA quantification by quantitative polymerase chain reaction and virus titration in Vero E6 cells. Viral RNA (Fig. 1, A and B) and infectious virus (Fig. 1, C and D) were detected in the nasal turbinate, soft palate, and tonsils of all four ferrets inoculated with these two viruses but were not detected in any other organs tested. These results indicate that SARS-CoV-2 can replicate in the upper respiratory tract of ferrets, but its replication in other organs is undetectable.

To investigate the replication dynamics of these virus strains in ferrets, groups of three animals were inoculated intranasally with 10^5 PFU of F13-E or CTan-H, and each ferret was then placed in a separate cage within an isolator. Nasal washes and rectal swabs were collected on days 2, 4, 6, 8, and 10 p.i. from the ferrets for viral RNA detection and virus titration. Body temperatures and signs of disease were monitored for 2 weeks. Viral RNA was detected in the nasal washes on days 2, 4, 6, and 8 p.i. in all six ferrets inoculated with the two strains (Fig. 1, E and F). Viral RNA was also detected in some of the rectal swabs of

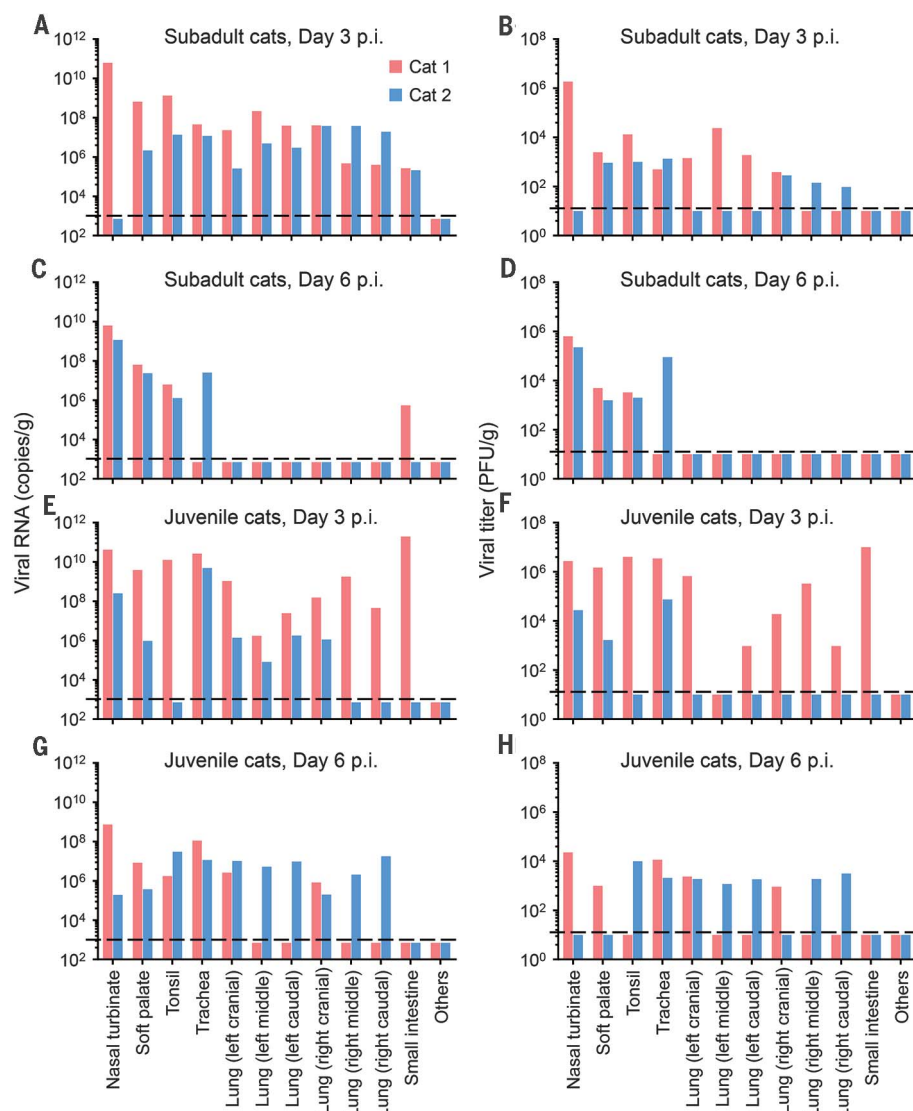


Fig. 2. Replication of SARS-CoV-2 in cats. Subadult cats and juvenile cats inoculated with CTan-H virus were euthanized on days 3 and 6 p.i., and their organs were collected for viral RNA detection and virus titration. (A) Viral RNA and (B) viral titers of subadult cats on day 3 p.i. (C) Viral RNA and (D) viral titers of subadult cats on day 6 p.i. (E) Viral RNA and (F) viral titers of juvenile cats on day 3 p.i. The values of red bars in (E) and (F) are from the cat that died on this day. (G) Viral RNA and (H) viral titers of juvenile cats on day 6 p.i. "Others" represents viral-negative organs, including the brain, heart, submaxillary lymph nodes, kidneys, spleen, liver, and pancreas. Each color bar represents the value from an individual animal. The dashed lines indicate the lower limit of detection.

the virus-inoculated ferrets, although the copy numbers were notably lower than those in the nasal washes of these ferrets (fig. S1, A and C). Infectious virus was detected from the nasal washes of all ferrets (Fig. 1, G and H) but not from the rectal swabs of any ferrets (fig. S1, B and D).

One ferret from each virus-inoculated group developed fever and loss of appetite on days 10 (CTan-H-inoculated) and 12 p.i. (F13-E-inoculated), respectively. To investigate whether these symptoms were caused by virus replication in the lower respiratory tract, we euthanized

the two ferrets on day 13 p.i. and collected their organs for viral RNA detection. However, viral RNA was not detected in any other tissues or organs of either ferret, except for a low copy number ($10^{5.4}$ copies/g) in the turbinate of the ferret inoculated with CTan-H (fig. S2). Pathological studies revealed severe lymphoplasmacytic perivascularitis and vasculitis; increased numbers of type II pneumocytes, macrophages, and neutrophils in the alveolar septa and alveolar lumen; and mild peribronchitis in the lungs of the two ferrets euthanized on day 13 p.i. (fig. S3). Antibodies

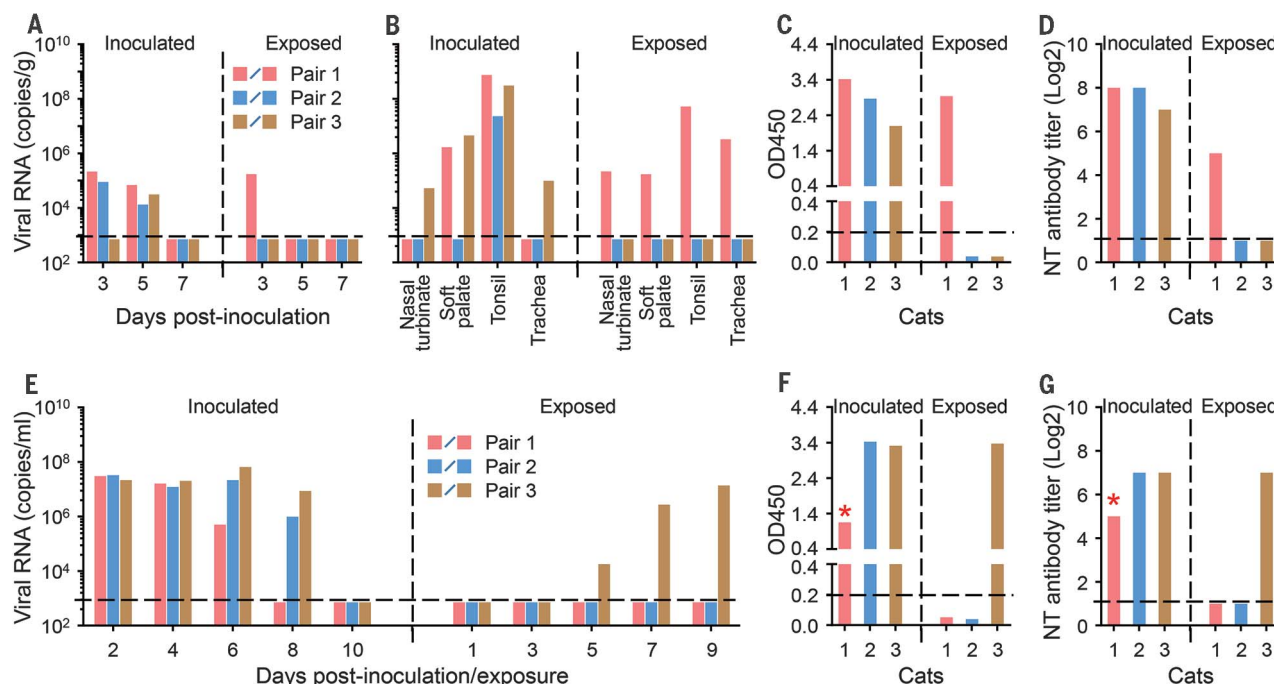


Fig. 3. Transmission of SARS-CoV-2 in cats. Transmission of the CTan-H virus strain was evaluated in subadult cats (A to D) and juvenile cats (E to G). (A) Viral RNA in the feces of virus-inoculated or virus-exposed subadult cats. (B) Viral RNA in tissues or organs of inoculated or exposed subadult cats euthanized on day 11 p.i. (pair 1, red bars) or day 12 p.i. (pairs 2 and 3). Antibodies against SARS-CoV-2 in these euthanized subadult cats were detected by ELISA (C) and neutralization assay (D). (E) Viral RNA in nasal washes of juvenile cats. Sera

from juvenile cats were collected on day 20 p.i., except for sera from one virus-inoculated animal that died on day 13 p.i. Antibody values for this cat (indicated by asterisks) were detected from sera collected on day 10 p.i.; antibodies against SARS-CoV-2 were detected by ELISA (F) and neutralization assay (G). Each color bar represents the value from an individual animal. The horizontal dashed lines in (C) and (F) show the cutoff value for seroconversion, and the horizontal dashed lines in the other panels indicate the lower limit of detection.

against SARS-CoV-2 were detected in all ferrets by an enzyme-linked immunosorbent assay (ELISA) and a neutralization assay, although the antibody titers of the two ferrets that were euthanized on day 13 p.i. were notably lower than those of the ferrets euthanized on day 20 p.i. (Fig. 1, I to L).

A virus attachment assay indicated that SARS-CoV-2 could attach to bronchiolar epithelial cells (fig. S4A) and some type II pneumocytes (fig. S4B) in ferret lungs. To further investigate whether SARS-CoV-2 replicates in the lungs of ferrets, we intratracheally inoculated eight ferrets with 10^5 PFU of CTan-H and euthanized two animals each on days 2, 4, 8, and 14 p.i. to look for viral RNA in the tissues and organs. Viral RNA was detected only in the nasal turbinate and soft palate of one ferret in each pair euthanized on days 2 and 4 p.i.; was detected in the soft palate of one ferret and in the nasal turbinate, soft palate, tonsils, and trachea of the other ferret euthanized on day 8 p.i.; and was not detected in either of the two ferrets euthanized on day 14 p.i. (fig. S5). These results indicate that SARS-CoV-2 can replicate in the upper respiratory tract of ferrets for up to 8 days without causing severe disease or death.

Cats and dogs are in close contact with humans; therefore, it is important to understand

their susceptibility to SARS-CoV-2. We first investigated the replication of SARS-CoV-2 in cats. Seven subadult cats (aged 6 to 9 months, outbred domestic cats) were intranasally inoculated with 10^5 PFU of CTan-H. Two animals were scheduled to be euthanized on days 3 and 6 p.i., respectively, to evaluate viral replication in their organs. Three subadult cats were placed in separate cages within an isolator. To monitor respiratory droplet transmission, an uninfected cat was placed in a cage adjacent to each of the infected cats. The aggressive behavior of the subadult cats made it difficult to perform regular nasal wash collection. To avoid possible injury, we only collected feces from these cats and checked for viral RNA in their organs after euthanasia.

Viral RNA was detected in the nasal turbinate of one animal, as well as in the soft palates, tonsils, tracheas, lungs, and small intestines of both animals euthanized on day 3 p.i. (Fig. 2A). In the animals euthanized on day 6 p.i., viral RNA was detected in the nasal turbinates, soft palates, and tonsils of both animals; in the trachea of one animal; and in the small intestine of the other. However, viral RNA was not detected in any lung samples from either of these animals (Fig. 2C). Infectious virus was detected in the viral RNA-positive nasal turbinates, soft palates, tonsils, tracheas, and

lungs of these cats but was not recovered from the viral RNA-positive small intestines (Fig. 2, B and D).

In the transmission study, viral RNA was detected in the feces of two virus-inoculated subadult cats on day 3 p.i. and in all three virus-inoculated subadult cats on day 5 p.i. (Fig. 3A). Viral RNA was detected in the feces of one exposed cat on day 3 p.i. (Fig. 3A). The pair of subadult cats with viral RNA-positive feces was euthanized on day 11 p.i. Viral RNA was detected in the soft palate and tonsils of the virus-inoculated animal and in the nasal turbinate, soft palate, tonsils, and trachea of the exposed animal (Fig. 3B), indicating that respiratory droplet transmission had occurred. We euthanized the other pairs on day 12 p.i. Viral RNA was detected in the tonsils of one virus-inoculated subadult cat and in the nasal turbinate, soft palate, tonsils, and trachea of the other virus-inoculated subadult cat but was not detected in any organs or tissues of the two exposed subadult cats (Fig. 3B). Antibodies against SARS-CoV-2 were detected in all three virus-inoculated subadult cats and one exposed cat via an ELISA and neutralization assay (Fig. 3, C and D).

We repeated the replication and transmission studies in juvenile cats (aged 70 to 100 days) (Figs. 2, E to H, and 3, E to G, and fig. S6).

Table 1. Susceptibility of dogs, pigs, chickens, and ducks to SARS-CoV-2. Animals were intranasally inoculated with 10⁵ PFU (dogs and pigs) or 10^{4.5} PFU (chickens and ducks) of the CTan-H virus strain. Two (dogs) or three (pigs, chickens, and ducks) uninfected animals were housed in the same room with their infected counterparts to monitor transmission of the virus. Oropharyngeal and rectal swabs from all animals were collected on the indicated days postinoculation (p.i.) for viral RNA detection. The “Other time points” category includes days 8, 10, 12, and 14 p.i.

		Viral RNA detection in animals inoculated with SARS-CoV-2 isolate CTan-H; positive cases/total (copies, log ₁₀)								Seroconversion; positive cases/total†
Animal	Treatment	Oropharyngeal swab				Rectal swab				
		Day 2 p.i.	Day 4 p.i.	Day 6 p.i.	Other time points	Day 2 p.i.	Day 4 p.i.	Day 6 p.i.	Other time points	
Dog*	Inoculated	0/5	0/5	0/4	0/4	2/5 (6.5, 5.4)	0/5	1/4 (4.2)	0/4	2/4
	Contact	0/2	0/2	0/2	0/2	0/2	0/2	0/2	0/2	0/2
Pig	Inoculated	0/5	0/5	0/5	0/5	0/5	0/5	0/5	0/5	0/5
	Contact	0/3	0/3	0/3	0/3	0/3	0/3	0/3	0/3	0/3
Chicken	Inoculated	0/5	0/5	0/5	0/5	0/5	0/5	0/5	0/5	0/5
	Contact	0/3	0/3	0/3	0/3	0/3	0/3	0/3	0/3	0/3
Duck	Inoculated	0/5	0/5	0/5	0/5	0/5	0/5	0/5	0/5	0/5
	Contact	0/3	0/3	0/3	0/3	0/3	0/3	0/3	0/3	0/3

*One virus-inoculated beagle was euthanized on day 4 p.i., but viral RNA was not detected in any of its collected organs, which included lung, trachea, nasal turbinate, soft palate, brain, heart, tonsils, kidneys, spleen, liver, pancreas, and small intestine (fig. S6). †Sera were collected from all animals on day 14 p.i., and antibodies against SARS-CoV-2 were detected by using a double-antigen sandwich ELISA kit (ProtTech, Luoyang, China).

Histopathologic studies performed on samples from the virus-inoculated juvenile cats that died or were euthanized on day 3 p.i. revealed massive lesions in the nasal and tracheal mucosa epitheliums and lungs (fig. S7). These results indicate that SARS-CoV-2 can replicate efficiently in cats and that younger cats are more vulnerable than older ones. Notably, our findings also reveal that the virus is transmissible between cats via the airborne route.

We next investigated the replication and transmission of SARS-CoV-2 in dogs. Five 3-month-old beagles were intranasally inoculated with 10⁵ PFU of CTan-H and housed with two uninoculated beagles in a room. Oropharyngeal and rectal swabs from each beagle were collected on days 2, 4, 6, 8, 10, 12, and 14 p.i. for viral RNA detection and virus titration in Vero E6 cells. Viral RNA was detected in the rectal swabs of two virus-inoculated dogs on day 2 p.i. and in the rectal swab of one such dog on day 6 p.i. (Table 1). One dog that was positive for viral RNA by rectal swab on day 2 p.i. was euthanized on day 4 p.i., but viral RNA was not detected in any organs or tissues collected from this animal (fig. S8). Additionally, infectious virus was not detected in any swabs collected from any of these dogs. Sera were collected from all dogs on day 14 p.i. for antibody detection by an ELISA. Two virus-inoculated dogs seroconverted; the other two virus-inoculated dogs and the two contact-exposed dogs were all seronegative for SARS-CoV-2 (Table 1 and fig. S9). These results indicate that dogs have low susceptibility to SARS-CoV-2.

We also investigated the susceptibility of pigs, chickens, and ducks to SARS-CoV-2 by

using the same strategy as that used to assess dogs. However, viral RNA was not detected in any swabs collected from these virus-inoculated animals or from naïve contact animals (Table 1). In addition, all of these animals were seronegative for SARS-CoV-2 when tested by ELISA with sera collected on day 14 p.i. (Table 1). These results indicate that pigs, chickens, and ducks are not susceptible to SARS-CoV-2.

In summary, we found that ferrets and cats are highly susceptible to SARS-CoV-2; dogs have low susceptibility; and pigs, chickens, and ducks are not susceptible to the virus. Unlike influenza viruses and the other SARS-coronavirus known to infect humans (SARS-CoV-1), which replicate in both the upper and lower respiratory tract of ferrets (20, 22–24, 26, 27), SARS-CoV-2 replicates only in the nasal turbinate, soft palate, and tonsils of ferrets. SARS-CoV-2 may also replicate in the digestive tract, as viral RNA was detected in the rectal swabs of the virus-infected ferrets, but virus was not detected in lung lobes, even after the ferrets were intratracheally inoculated with the virus. It remains unclear whether the virus causes more severe disease in male ferrets than in female ferrets, as has been observed among humans (13, 28).

Several studies have reported that SARS-CoV-2 uses angiotensin-converting enzyme 2 (ACE2) as its receptor to enter cells (3, 29–31). ACE2 is mainly expressed in type II pneumocytes and serous epithelial cells of tracheobronchial submucosal glands in ferrets (25). Ferrets and cats differ by only two amino acids in the SARS-CoV-2 spike-contacting regions of ACE2 (table S1); therefore, the underlying mechanism that prevents the replication of

SARS-CoV-2 in the lower respiratory tract of ferrets remains to be investigated. The fact that SARS-CoV-2 replicates efficiently in the upper respiratory tract of ferrets makes them a candidate animal model for evaluating the efficacy of antiviral drugs or vaccines against COVID-19.

The cats we used in this study were outbred and were susceptible to SARS-CoV-2, which replicated efficiently and was transmissible to naïve cats. Cats in Wuhan have been reported to be seropositive for SARS-CoV-2 (32). Surveillance for SARS-CoV-2 in cats should be considered as an adjunct to elimination of COVID-19 in humans.

REFERENCES AND NOTES

1. L. Zou et al., *N. Engl. J. Med.* **382**, 1177–1179 (2020).
2. N. Zhu et al., *N. Engl. J. Med.* **382**, 727–733 (2020).
3. P. Zhou et al., *Nature* **579**, 270–273 (2020).
4. F. Wu et al., *Nature* **579**, 265–269 (2020).
5. A. Wu et al., *Cell Host Microbe* **27**, 325–328 (2020).
6. C. Wang, P. W. Horby, F. G. Hayden, G. F. Gao, *Lancet* **395**, 470–473 (2020).
7. Y. Pan, D. Zhang, P. Yang, L. L. M. Poon, Q. Wang, *Lancet Infect. Dis.* **20**, 411–412 (2020).
8. X. Pan et al., *Lancet Infect. Dis.* **20**, 410–411 (2020).
9. R. Lu et al., *Lancet* **395**, 565–574 (2020).
10. P. Liu, X. Z. Tan, *Radiology* **295**, 19 (2020).
11. Q. Li et al., *N. Engl. J. Med.* **382**, 1199–1207 (2020).
12. C. Huang et al., *Lancet* **395**, 497–506 (2020).
13. W. J. Guan et al., *N. Engl. J. Med.* 10.1056/NEJMoa2002032 (2020).
14. N. Chen et al., *Lancet* **395**, 507–513 (2020).
15. J. F. Chan et al., *Lancet* **395**, 514–523 (2020).
16. J. F. Chan et al., *Emerg. Microbes Infect.* **9**, 221–236 (2020).
17. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, *Nat. Microbiol.* **5**, 536–544 (2020).
18. World Health Organization (WHO), “Coronavirus disease 2019 (COVID-19): Situation Report – 51” (WHO, 2020); www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf?sfvrsn=1ba62e57_4.
19. See supplementary materials.
20. J. Shi et al., *Cell Res.* **27**, 1409–1421 (2017).
21. K. J. Stittelaar et al., *Viruses* **8**, 168 (2016).

22. Q. Zhang *et al.*, *Science* **341**, 410–414 (2013).
23. M. Imai *et al.*, *Nature* **486**, 420–428 (2012).
24. S. Herfst *et al.*, *Science* **336**, 1534–1541 (2012).
25. J. M. van den Brand *et al.*, *Vet. Pathol.* **45**, 551–562 (2008).
26. B. E. Martina *et al.*, *Nature* **425**, 915 (2003).
27. Y. K. Chu *et al.*, *Virology* **374**, 151–163 (2008).
28. Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, *Zhonghua Liu Xing Bing Xue Za Zhi* **41**, 145–151 (2020) [in Chinese].
29. R. Yan *et al.*, *Science* **367**, 1444–1448 (2020).
30. M. Letko, A. Marzi, V. Munster, *Nat. Microbiol.* **5**, 562–569 (2020).
31. The *Cell* Editorial Team, *Cell* **181**, 1–3 (2020).
32. Q. Zhang *et al.*, SARS-CoV-2 neutralizing serum antibodies in cats: A serological investigation. *bioRxiv* (2020). <https://doi.org/10.1101/2020.04.01.021196>.

ACKNOWLEDGMENTS

We thank S. Watson for editing the manuscript. **Funding:** This work was supported by the National Key R&D Program of China

(2020YFC0846500, 2018YFC1200601, and 2016YFD0500301).

Author contributions: J.S., Z.W., G.Z., H.Y., C.W., B.H., R.L., X.H., L.S., Z.S., Y.Z., P.L., L.L., P.C., J.W., X.Z., and Y.G. performed experiments; J.S., Z.W., G.Z., H.Y., C.W., W.T., G.W., H.C., and Z.B. analyzed data; and Z.B. and H.C. designed the study and wrote the manuscript. **Competing interests:** None of the authors have any competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. Sequences of the viruses used in this study have been deposited in GISAID previously with the accession numbers EPI_ISL_402119 and EPI_ISL_408514. Two strains of 2019 novel coronavirus (CTan-H and F13-E) were obtained from the China CDC under a material transfer agreement that allows use only in P3+ or P4 facilities, prevents live virus sharing, and prevents commercial use. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>.

This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6494/1016/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S10
Table S1
References and Notes (33–35)
MDAR Reproducibility Checklist

[View/request a protocol for this paper from Bio-protocol.](#)

12 March 2020; accepted 7 April 2020
Published online 8 April 2020
10.1126/science.abb7015

ICE SHEETS

Delicate seafloor landforms reveal past Antarctic grounding-line retreat of kilometers per year

J. A. Dowdeswell^{1*}, C. L. Batchelor^{1,2}, A. Montelli¹, D. Ottesen³, F. D. W. Christie¹, E. K. Dowdeswell¹, J. Evans⁴

A suite of grounding-line landforms on the Antarctic seafloor, imaged at submeter horizontal resolution from an autonomous underwater vehicle, enables calculation of ice sheet retreat rates from a complex of grounding-zone wedges on the Larsen continental shelf, western Weddell Sea. The landforms are delicate sets of up to 90 ridges, <1.5 meters high and spaced 20 to 25 meters apart. We interpret these ridges as the product of squeezing up of soft sediment during the rise and fall of the retreating ice sheet grounding line during successive tidal cycles. Grounding-line retreat rates of 40 to 50 meters per day (>10 kilometers per year) are inferred during regional deglaciation of the Larsen shelf. If repeated today, such rapid mass loss to the ocean would have clear implications for increasing the rate of global sea level rise.

The ice shelves fringing about 75% of the Antarctic Ice Sheet represent a highly sensitive interface between ice and ocean, with potential for rapid grounding-line retreat and associated mass loss from the parent ice sheet (1, 2). It is not known, however, whether modern satellite-derived grounding-line retreat rates of tens to hundreds of meters per year, acquired over a time window of about 30 years at most (3, 4), are representative of the maximum possible magnitude of retreat in, for example, West Antarctica's Pine Island Bay.

An ice sheet grounding zone is the region over which seaward-flowing ice decouples from its underlying bed and becomes a freely floating ice shelf; the instantaneous, tidally modulated junction between grounded ice, the seafloor, and the resulting sub-ice shelf ocean-water cavity beyond is known as the grounding line (5). The former grounding zone is often identified on high-latitude continental shelves by the presence of sedimentary depositional centers, or grounding-zone wedges (GZWs), which provide a well-preserved geological archive of the transition zone between grounded ice sheets and floating ice shelves (5–7). GZWs build up predominantly through delivery of soft deforming sediments from the ice sheet interior to the grounding zone along a line source. There is little space for vertical accretion in the restricted cavities immediately beyond the grounding zone (8), and GZWs are therefore typically subdued, asymmetrical features on the seafloor; they generally have a steeper ice-distal face of a few degrees in slope angle and a more extensive ice-proximal portion often less than 1° (5, 7). GZW volumes can

be up to several cubic kilometers, with the rate of formation depending on sediment supply and the duration of any still-stand during regional grounding-line retreat (5, 9).

We investigated, using an autonomous underwater vehicle (AUV) (10), the morphology and shallow stratigraphy of an ~9-km² area of five GZWs within a 40-km-by-10-km grounding-zone complex preserved after ice retreat beneath about 500 m of water on the

continental shelf offshore of Larsen Inlet, eastern Antarctic Peninsula (Figs. 1 and 2 and fig. S1) (11). Multibeam echo-sounder data from the AUV have a horizontal resolution of a few tens of centimeters (materials and methods), representing a step-change in our ability to observe fine-scale landforms preserved on the seafloor and to understand the highly sensitive setting of a former ice sheet grounding zone and the processes operating there. The GZWs were deposited during deglaciation of the northeastern Antarctic Peninsula continental shelf following the Last Glacial Maximum (LGM), probably before the minimum age for the transition from grounded ice to ice shelf in the inner Larsen A embayment at 10,700 calibrated years before the present [derived by relative paleomagnetic intensity dating of core KC023, Fig. 1A (11–13)]. They have remained almost undisturbed since ice retreat, except for a thin drape of hemipelagic sediment (Fig. 2).

Surface-morphological evidence (Fig. 2A) and shallow acoustic stratigraphy (Fig. 2, C and D) demonstrate that the GZWs cross-cut one another, providing a relative chronology for formation (inset of Fig. 2B). Heights varying between about 10 and 20 m mark the outer edges of GZWs 3 to 5 and, together with limited acoustic-stratigraphic evidence (Fig. 2, C and D), imply that the depocenters are about

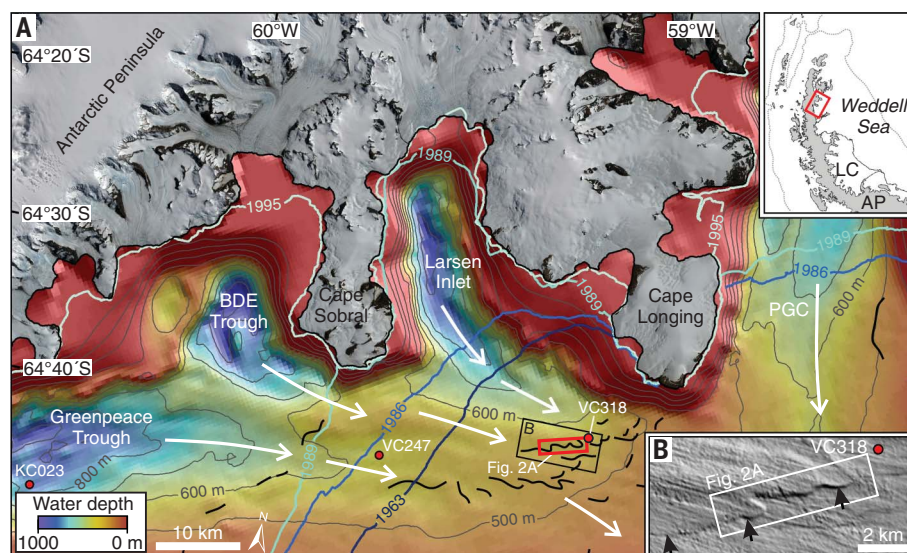


Fig. 1. Map of the study area. (A) Map showing the location of the surveyed GZW complex in Larsen Inlet (red box). Black lines are GZW crests (11). White arrows show the former ice flow direction. Red circles are locations of sediment cores KC023, VC247, and VC318. Background land area is from a Sentinel-2b Level-1C Top of Atmosphere reflectance image acquired on 5 December 2018. Background bathymetry is from the International Bathymetric Chart of the Southern Ocean (IBCSO) (38). BDE Trough is Bombardier, Dinsmoor and Edgemoor Trough. The 1986, 1989, and 1995 ice shelf frontal positions are from Cook *et al.* (39), and the 1963 ice shelf front position was ascertained from declassified Argon satellite imagery (40). The inset shows the location of the study area (red box) on the Antarctic Peninsula (AP). The dashed gray line is the present-day continental shelf break. LC, Larsen C Ice Shelf. **(B)** Bathymetric image of the GZW complex beyond Larsen Inlet, acquired from the RRS *James Clark Ross* in 2002 using a Kongsberg EM120 multibeam echo sounder with a frequency of 12 kHz. Grid-cell size is 50 m. The image is modified from Evans *et al.* (11). Black arrows show GZW crests.

¹Scott Polar Research Institute, University of Cambridge, Cambridge, UK. ²Norwegian University of Science and Technology, Trondheim, Norway. ³Geological Survey of Norway, Trondheim, Norway. ⁴Department of Geography, Loughborough University, Loughborough, UK.

*Corresponding author. Email: jd16@cam.ac.uk

10 to 20 m thick. Their asymmetry (Fig. 2, A and D) is typical of more than 100 GZWs identified elsewhere in the polar continental-shelf record (5, 7). GZWs “a” and 2 to 5 rest on an older surface (labeled 1 in Fig. 2B) that includes streamlined landforms (Fig. 2A), and nearby cores VC318 and VC247 (Fig. 1A) contain soft (<8 kPa) diamictic subglacial traction till that has undergone deformation (11, 14). These elongate sedimentary features, which are interpreted as mega-scale glacial lineations (MSGs) (Fig. 2, A, E, and F), are orientated in the direction of regional ice flow (Fig. 1A). The presence of MSGs implies formation subglacially in soft sediments beneath actively flowing and probably fast-flowing ice before and during GZW deposition (15, 16). In addition, the ice-distal margin of the GZWs, marking former ice sheet grounding zones, contains a number of lobate forms 50 to 100 m long (Fig. 2, E and F) that are interpreted as debris flows, demonstrating wedge progradation as deformation till was delivered to the grounding line by active ice.

On the low-gradient GZW surfaces, a complex assemblage of fine-scale landforms is present, imaged at submeter resolution (Fig. 2A). The features are of two types, forming the appearance of “ladders” with numerous “rungs.” The sides of each ladder are linear features typically 2 to 4 m high, hundreds of meters long, spaced 50 to 200 m apart, and orientated subparallel to the past ice flow direction; they are interpreted as MSGs. The MSGs are formed within an acoustically semitransparent unit that is overlain by a thin (about 1 m) drape of sediment (Fig. 2, C and D). Overprinting the MSGs are sets of delicate transverse-to-flow ridges or rungs that are typically <0.5 m high and spaced about 20 to 25 m apart (Table 1). These rungs are of relatively uniform morphology, with many having a steeper ice-distal side (Fig. 3, B to E).

This delicate submarine glacial landform assemblage, which we term “ladder and rung topography,” covers almost the whole imaged area of each GZW (Figs. 2A and 3). A maximum of 90 rungs have been identified on GZW 4 (Fig. 3 and Table 1). Although the rungs have greater seafloor expression in the depressions between MSGL crests, they can often be traced across the intervening and larger MSGs, and series of up to 50 are continuous right across several of the GZWs (Fig. 3A). This implies that the rungs are younger than the larger flow-parallel MSGs and the GZWs. The rungs are interpreted to have formed during retreat of the ice sheet grounding line across the GZW surface. Their regular spacing implies a cyclical formation mechanism.

Regularly spaced small transverse ridges have been observed occasionally at coarser resolution on high-latitude continental shelves and in fjords (17, 18). Graham *et al.* (18), for

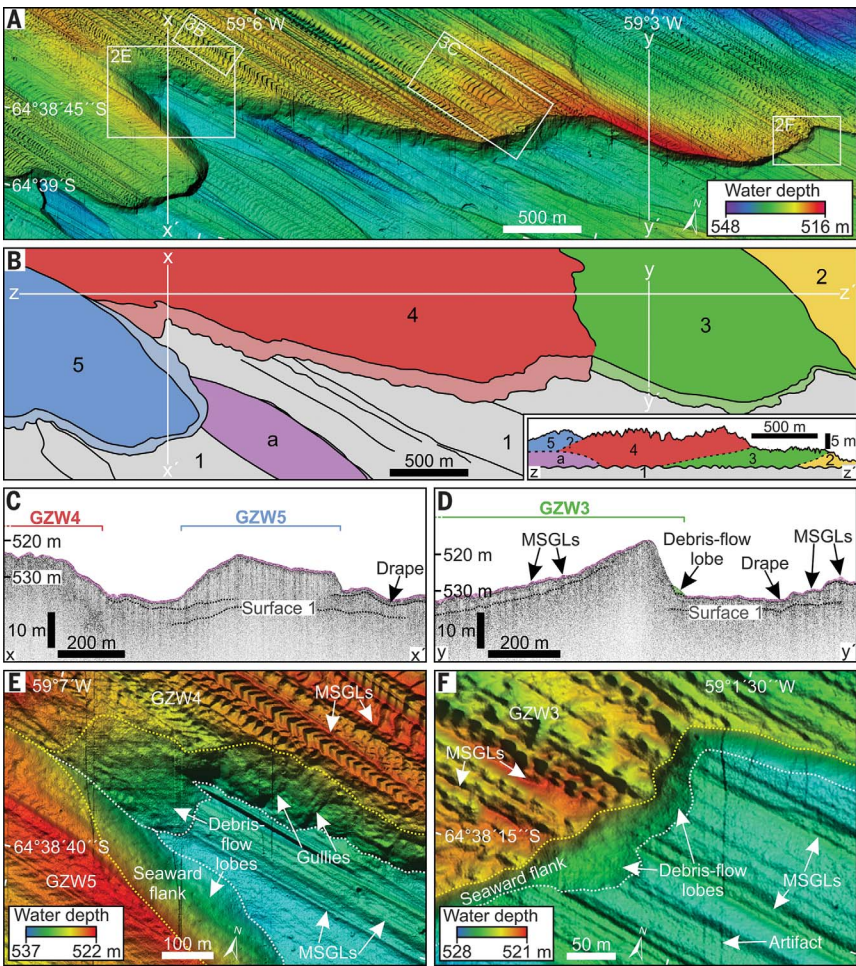


Fig. 2. Geophysical data showing a GZW complex in Larsen Inlet, acquired from an AUV. (A) Bathymetric data of the GZW complex, derived from an AUV-deployed multibeam echo sounder. Grid-cell size is 1 m. The location is shown in Fig. 1A. The boxes labeled 2E, 3B, 3C, and 2F show the locations of areas detailed in the corresponding figure panels. **(B)** Interpretation of individual GZWs within the grounding-zone complex. The inset is a schematic diagram showing GZW stratigraphy as derived from multibeam echo-sounder data and sub-bottom profiles. GZWs “a” and 2 to 5 rest on an older surface (labeled 1). GZW “a” overlies surface 1 and is beneath GZW 5, but its stratigraphic position relative to GZWs 2 to 4 is unknown. **(C and D)** Sub-bottom profiles along the GZWs, showing the GZW stratigraphy and a thin (~0.7 m) uppermost draping unit (pink fill). **(E and F)** Detail of debris-flow lobes on the seaward flank of the GZWs.

Table 1. Morphological characteristics of rungs or transverse-to-flow ridges on GZW surfaces in the study area. GZW thickness is maximum thickness derived from sub-bottom profiles. Dashed lines indicate where GZW thickness was not able to be ascertained. GZW area is only given for those parts of the GZWs that were surveyed from the AUV (Fig. 2A); the GZWs most likely extend beyond the study area. Average daily retreat and retreat duration values are calculated assuming that the ridges were produced by tidally influenced sit-downs of the grounding line (i.e., two rungs or ridges produced per day).							
GZW surface	GZW maximum thickness (m)	GZW planar area (km ²)	Number of ridges	Ridge height (m)	Average ridge spacing (m)	Average daily retreat (m)	Retreat duration (days)
1 (several)	—	3.1	71	0.1–0.4	20	40	35
a	3	0.4	33	0.2–0.3	24	48	16
2	—	0.3	28	0.2–0.6	25	50	14
3	16	1.4	44	0.3–1.0	20	40	22
4	15	2.3	90	0.5–1.5	21	42	45
5	9	1.0	40	0.2–0.4	20	40	20

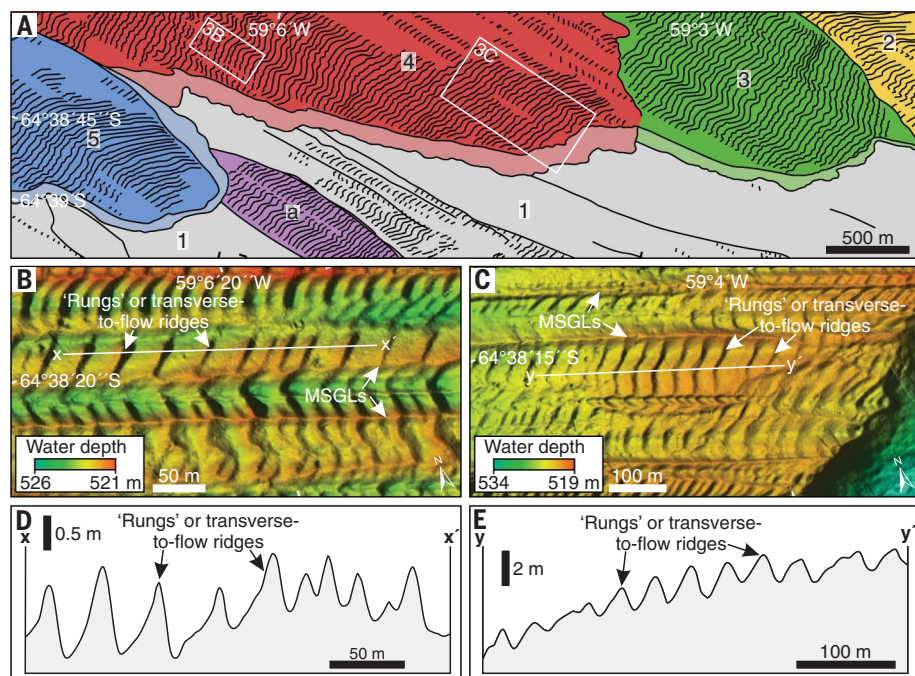


Fig. 3. Detail of ladder and rung topography. (A) The mapped distribution of rungs or transverse-to-flow ridges on GZW surfaces. The boxes labeled 3B and 3C show the locations of areas detailed in (B) and (C). (B and C) Details of transverse-to-flow ridges on GZW surfaces. (D and E) Profiles across the rungs shown in (B) and (C).

example, used an AUV to image distinctive sets of ridges along several individual transects beneath the floating tongue of Pine Island Glacier, West Antarctica, at 2-m grid-cell size. The dimensions, spacing, and plan-view and cross-sectional geometry of similar ridges observed on several Antarctic shelves are compared in table S1. Many of the sets of small ridges, which have been referred to previously as corrugation ridges, have been interpreted as the product of tidal action that moves ice regularly up and down on a sedimentary seafloor (19). The ice can be in the form of iceberg keels, sub-ice shelf keels, or a wider grounding line (supplementary text and table S1). Where corrugation ridges are located within iceberg ploughmarks, they are likely to have been produced by the tidal motion of iceberg keels impinging on the seafloor (17). Corrugation ridges are also present within scours that are interpreted to have been produced by the forward ploughing of ice keels close to the grounding line (18). Where these ridges, or rungs, are more laterally extensive, as we observe over several kilometers on the Larsen shelf, and where they overprint MSGLs on the surface of GZWs (Fig. 2A), they likely form through regular vertical motion of an ice sheet grounding line that leads to the squeezing up of deforming soft sediment as small ridges on each falling tide (18). Indeed, it has been observed that short-lived perturbations in upstream longitudinal stress and ice flow, associated with tidal-induced flexure

at the grounding zone, can cause small changes in the surface of the Antarctic Ice Sheet even tens of kilometers inland of the grounding line (20, 21). By contrast, recessional moraine ridges, which are typically larger and less regularly spaced than the rungs reported here (supplementary text and table S1), have been described from the Ross and Amundsen seas and are thought to form by short-lived readvances of the grounding line during overall retreat (22–24). In Svalbard fjords, such ridges form during small winter readvances of tidewater glacier termini during regional glacier recession (25, 26).

Given their presence overprinting MSGLs on GZW surfaces, and continuity over at least 5 km, the rungs of our ladder and rung topography are interpreted to form by grounding-line sediment squeezing in successive tidal cycles (Figs. 2 and 3). Grounding-line retreat is necessary to preserve individual rungs produced as the ice presses into the underlying soft sediment on each falling tide; otherwise, the delicate rung would be deformed or partially eroded during the next tidal cycle. The individual rungs are generally well developed and simple in morphology, with little evidence of subsequent disturbance (Figs. 2 and 3). We provide a schematic model of the formation and preservation of ladder and rung topography at the retreating grounding line of an ice shelf (Fig. 4). Although we favor an interpretation in which rung formation occurs at the

grounding line, it is possible that these features were produced subglacially, contemporaneous with the MSGLs. Subglacial processes for the formation of subdued transverse ridges could include sediment squeezing into basal crevasses as the ice moves forward and/or water flow over sediment in the depressions between the MSGLs (18).

Our interpretation of ladder and rung topography as a product of rapid, tidally modulated grounding-line migration allows the past rate of deglacial retreat to be calculated over days to weeks. This is only possible due to the sub-meter resolution of our AUV-derived multi-beam imagery (Figs. 2 and 3). In the absence of systematic tidal-gauge observations, an inverse model of Weddell Sea tides by Padman *et al.* (27) predicts semidiurnal tidal heights generally <1.5 m for Larsen area today; maximum tidal ranges can exceed 3 m in some places. During ship operations adjacent to Larsen C Ice Shelf in January 2019, we confirmed a strong coherency between the predictions of Padman *et al.*'s tidal model and our onboard observations. Although the configuration of the Weddell Sea during deglaciation will differ in detail from today owing to isostatic rebound and the presence of residual inner-shelf ice, gross basin form will be largely similar.

The rate of deglacial grounding-line retreat is derived from the spacing between each rung. Such an implication would be valid whether the rungs were produced along a whole section of the grounding line or by discrete ice-shelf keels impinging on the seafloor (18), although we favor the former mechanism for the Larsen shelf features given their considerable lateral continuity. Average rung spacing is between 20 and 25 m, and the number of observed rungs in the imaged area of each GZW is between 28 and 90 (Table 1); there are likely to be many more rungs on the unimaged distal surfaces of the GZWs. Assuming formation during each semidiurnal tidal cycle, this yields an average grounding-line retreat between 40 and 50 m/day over periods between 14 and 45 days. If extrapolated over the GZW surfaces landward of our imaged area, this gives a grounding-line retreat of about 18 km/year, which might conservatively be halved to about 10 km if we assume that winter sea ice cover might curtail ice shelf flexure associated with long-period swell waves for about half the year (28). This retreat rate of many kilometers per year is at least an order of magnitude higher than that observed between 1992 and 2011 at Pine Island Glacier (1.6 km/year) (29), which was itself two orders of magnitude greater than the average retreat rate for the past 10,000 years (30). Short-lived phases of very rapid post-LGM retreat, of up to 100 km/year, have also been simulated in numerical experiments by Jamieson *et al.* (31). The alternative model of rung formation, in

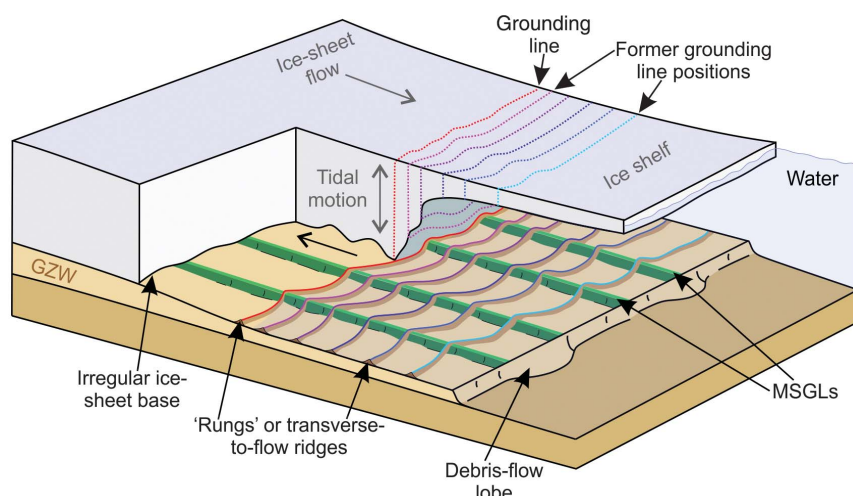


Fig. 4. Schematic model (not to scale) of the hypothesized formation of rungs or transverse-to-flow ridges on the surface of a GZW. When a glacier transitions into a floating ice shelf, delicate transverse-to-flow ridges can form by tidally influenced “sit-downs” of the grounding line. The ridges are formed at low tide by ice squeezing and pushing of sediments at the grounding line. The plan-view shape of the ridges reflects lateral variations in the shape of the grounding line. Series of parallel to subparallel ridges are formed when there is continuous grounding-line retreat. Because the ice is actively flowing during grounding-line retreat, as indicated by the presence of MSGLs, the parallel to subparallel nature of the rungs or ridges suggests that the shape of the ice sheet base remained relatively constant in the ice flow direction.

which the rungs are produced subglacially (18), implies even higher rates of grounding-line retreat, because the ice would have to lift off from its bed near-instantaneously to preserve the delicate rungs.

The implication is that retreat across the imaged area of the GZW complex on the Larsen continental shelf could have taken place in about 1 year. The multiple sets of GZWs suggest that periods of readvance and still-stand punctuated overall retreat, demonstrating the highly dynamic behavior of the grounding line. Presumably, final rapid retreat would have exposed the distal sections of the GZW complex and continued through the remaining 30 km or so of Larsen Inlet, given the deepening water landward of our study area (Fig. 1A) and the increasing buoyancy of the grounding zone that would result. Thus, it appears that final deglacial retreat from the still-stand represented by the last of the 10- to 20-m-thick GZWs was very rapid. This explains, too, why many streamlined glacial landforms, produced subglacially before deglaciation, are typically well preserved and not overprinted by other landforms—retreat is sufficiently rapid that there is little deposition. Grounding-line retreat, as inferred here over each tidal cycle, is not necessarily accompanied by retreat of the ice shelf frontal margin. For example, Pine Island Glacier underwent a sustained grounding-line retreat of ~30 km between 1992 and 2011 (3, 4) but did not experience any substantial change in its ice-shelf frontal position during that time (32).

The wider importance of our high-resolution observations of a past ice shelf GZW complex is to show, using the geological record, that very high rates of grounding-line retreat are possible, greater by an order of magnitude than those reported for ice shelves since satellite observations began. Thus, a grounding-line retreat of, for example, 10 km/year along a 5 km length of a 500-m-thick ice stream would yield a total ice mass loss of 25 km³/year [23 billion tonnes (Gt)/year using an assumed ice density of 916.7 kg/m³]. A similar calculation extrapolated across a 30-km-wide and 1-km-thick grounding line, the typical width and thickness of most ice stream centers in modern Pine Island Bay in West Antarctica, yields a total mass loss of 150 km³/year (138 Gt/year). This value is three to five times greater than mean contemporary (1992–2017) rates of mass loss integrated over Pine Island Glacier’s entire drainage basin (31 km³/year; 28.4 Gt/year) and the combined Thwaites–Pope–Smith–Kohler glacier drainage system (50.3 km³/year; 46.1 Gt/year), respectively (33). This implies that the rapid retreat of even a single Antarctic outlet glacier could, therefore, substantially increase the short-term mass loss from the ice sheet. This is without taking account of residual ice-dynamic effects associated with ice shelf loss (34); the removal of this buttressing effect provides an additional mechanism for substantial increases in ice discharge to the ocean from interior drainage basins (1, 2, 35). In addition, once retreat at this pace has begun, the self-stabilizing process of continuing subglacial

sediment delivery to the grounding line to offset deglacial ice shelf thinning and/or sea level rise (36, 37) ceases to be important given the very short time, perhaps only a year or two, for sediment buildup.

REFERENCES AND NOTES

1. T. K. Dupont, R. B. Alley, *Geophys. Res. Lett.* **32**, L04503 (2005).
2. D. Goldberg, D. M. Holland, C. Schoof, *J. Geophys. Res. Earth Surf.* **114**, F04026 (2009).
3. J. W. Park et al., *Geophys. Res. Lett.* **40**, 2137–2142 (2013).
4. E. Rignot, J. Mouginot, B. Scheuchl, MEaSUREs Antarctic grounding line from differential satellite radar interferometry, version 2. NASA National Snow and Ice Data Center Distributed Active Archive Center (2016); doi: 10.5067/IKBWW4RYHF1Q.
5. C. L. Batchelor, J. A. Dowdeswell, *Mar. Geol.* **363**, 65–92 (2015).
6. R. B. Alley, D. D. Blankenship, S. T. Rooney, C. R. Bentley, *Mar. Geol.* **85**, 101–120 (1989).
7. J. A. Dowdeswell, E. M. G. Fugelli, *Bull. Geol. Soc. Am.* **124**, 1750–1761 (2012).
8. K. Christianson et al., *J. Geophys. Res.* **121**, 1954–1983 (2016).
9. J. A. Dowdeswell et al., in *Atlas of Submarine Glacial Landforms: Modern, Quaternary and Ancient*, J. A. Dowdeswell et al., Eds. (Geological Society, London, Memoirs Series, 2016), vol. 46, pp. 519–552.
10. J. A. Dowdeswell et al., *J. Glaciol.* **54**, 661–672 (2008).
11. J. Evans, C. J. Pudsey, C. Ó Cofaigh, P. W. Morris, E. W. Domack, *Quat. Sci. Rev.* **24**, 741–774 (2005).
12. S. Brachfeld et al., *Geology* **31**, 749–752 (2003).
13. C. Ó Cofaigh et al., *Quat. Sci. Rev.* **100**, 87–110 (2014).
14. C. J. Pudsey, J. W. Murray, P. Appleby, J. Evans, *Quat. Sci. Rev.* **25**, 2357–2379 (2006).
15. C. D. Clark, *Earth Surf. Process. Landf.* **18**, 1–29 (1993).
16. E. C. King, R. C. A. Hindmarsh, C. R. Stokes, *Nat. Geosci.* **2**, 585–588 (2009).
17. M. Jakobsson et al., *Geology* **39**, 691–694 (2011).
18. A. G. C. Graham et al., *J. Geophys. Res.* **118**, 1356–1366 (2013).
19. J. A. Smith et al., *Nat. Commun.* **10**, 5635 (2019).
20. R. A. Bindenschadler, M. A. King, R. B. Alley, S. Anandakrishnan, L. Padman, *Science* **301**, 1087–1089 (2003).
21. B. M. Minchew, M. Simons, B. Riel, P. Milillo, *J. Geophys. Res.* **122**, 167–190 (2017).
22. A. R. W. Halberstadt, L. M. Simkins, S. L. Greenwood, J. B. Anderson, *Cryosphere* **10**, 1003–1020 (2016).
23. J. P. Klages et al., in *Atlas of Submarine Glacial Landforms: Modern, Quaternary and Ancient*, J. A. Dowdeswell et al., Eds. (Geological Society, London, Memoirs Series, 2016), vol. 46, pp. 349–352.
24. L. M. Simkins et al., *Nat. Geosci.* **10**, 691–697 (2017).
25. G. S. Boulton, *Sedimentology* **33**, 677–698 (1986).
26. D. Ottesen, J. A. Dowdeswell, *J. Geophys. Res.* **111**, F01016 (2006).
27. L. Padman, H. A. Fricker, R. Coleman, S. Howard, L. Erofeeva, *Ann. Glaciol.* **34**, 247–254 (2002).
28. R. A. Massom et al., *Nature* **558**, 383–389 (2018).
29. E. Rignot, J. Mouginot, M. Morlighem, H. Seroussi, B. Scheuchl, *Geophys. Res. Lett.* **41**, 3502–3509 (2014).
30. C.-D. Hillenbrand et al., *Geology* **41**, 35–38 (2013).
31. S. S. R. Jamieson et al., *Nat. Geosci.* **5**, 799–802 (2012).
32. H. D. Pritchard et al., *Nature* **484**, 502–505 (2012).
33. A. Shepherd et al., *Geophys. Res. Lett.* **46**, 8174–8183 (2019).
34. J. Mouginot, E. Rignot, B. Scheuchl, *Geophys. Res. Lett.* **41**, 1576–1584 (2014).
35. T. A. Scambos, J. A. Bohlander, C. A. Shuman, P. Skvarca, *Geophys. Res. Lett.* **31**, L18402 (2004).
36. R. B. Alley, S. Anandakrishnan, T. K. Dupont, B. R. Parizek, D. Pollard, *Science* **315**, 1838–1841 (2007).
37. S. Anandakrishnan, G. A. Catania, R. B. Alley, H. J. Horgan, *Science* **315**, 1835–1838 (2007).
38. J. E. Arndt et al., *Geophys. Res. Lett.* **40**, 3111–3117 (2013).
39. A. J. Cook, D. G. Vaughan, A. J. Luckman, T. Murray, *Antarct. Sci.* **26**, 614–624 (2014).
40. K. Kim, K. C. Jezek, H. Liu, *Int. J. Remote Sens.* **28**, 5357–5373 (2007).

ACKNOWLEDGMENTS

AUV deployments were from the South African icebreaking research vessel S.A. Agulhas II, and we thank the captain, ice pilot, officers, and crew for their support of our work. The Copernicus/ESA Sentinel-2b image shown in Fig. 1A was acquired free-of-charge from the Copernicus Open Access Hub (<https://scihub.copernicus.eu/>). **Funding:** This research was

funded by the Flotilla Foundation and Marine Archaeology Consultants Switzerland. The AUV and associated technical support were provided by Ocean Infinity and Deep Ocean Search. During this work, C.L.B. was in receipt of a grant from the Norwegian VISTA program. **Author contributions:** All authors took part in the acquisition and interpretation of the marine-geophysical data included in this paper. J.A.D. wrote the paper, and each co-author contributed comments and

suggestions during a number of iterations of the text and figures. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are shown in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6494/1020/suppl/DC1

Materials and Methods
Supplementary Text
Fig. S1
Table S1
References (41–57)

28 August 2019; accepted 16 April 2020
10.1126/science.aaz3059

10 ways that Science Careers can help advance your career

1. Register for a free online account on **ScienceCareers.org**.
2. Search thousands of job postings and find your perfect job.
3. Sign up to receive e-mail alerts about job postings that match your criteria.
4. Upload your resume into our database and connect with employers.
5. Watch one of our many webinars on different career topics such as job searching, networking, and more.
6. Download our career booklets, including Career Basics, Careers Beyond the Bench, and Developing Your Skills.
7. Complete an interactive, personalized career plan at “my IDP.”
8. Visit our Employer Profiles to learn more about prospective employers.
9. Research graduate program information and find a program right for you.
10. Read relevant career advice articles from our library of thousands.

Visit **ScienceCareers.org**
today — all resources are free



Science Careers

FROM THE JOURNAL SCIENCE  AAAS

SCIENCECAREERS.ORG

By Tess Torregrosa

You never know until you try

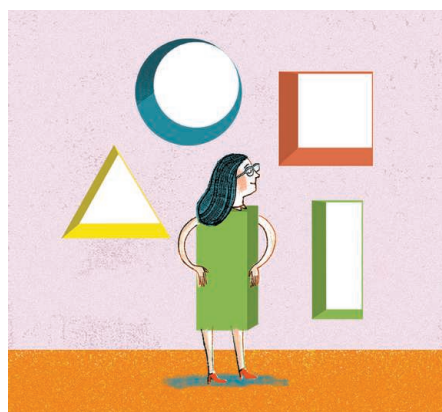
One hot July morning, I put on a business casual dress and headed to a biotech company for the first day of my internship. It felt daunting. Academia was my safe space; I had never worked anywhere else. When I arrived at the office, I was ushered into an orientation room, where people in formal business attire told me about the company's mission, products, and research areas. As the morning wore on, my unease about working for a for-profit company grew. I left the orientation wondering whether pausing my Ph.D. research to take on the internship had been a mistake.

When I started my Ph.D. program, I was open with my advisers about my career intentions, telling them that I wanted to work outside academia. I enjoyed doing research, but I wasn't interested in sitting at a desk writing grants for the rest of my career—something I knew I'd have to do a lot of if I became a professor.

My university had a program that enabled students to take a break from their research to do an internship at a company or government agency. My advisers told me about the program early on, and they encouraged me to plan on doing an internship during my final year. They thought that the experience would help me build connections and figure out what to do after graduation.

So I started to search for internships during my fourth year, at first focusing my search on science policy because I'd grown increasingly interested in that career path. I was disheartened to find that paid internships in science policy in my city were few and far between. I needed a solid paycheck and couldn't afford to relocate to another city temporarily, so I widened my search and started to look at positions in the biotech industry. My university had connections with local companies, so it was easier to find industry positions.

After the orientation session on my first day at the company, my manager sat me down to chat. I noticed her sneakers, colorful T-shirt, and cool haircut, and I remember thinking that she didn't look anything like the corporate representatives I'd listened to that morning. She asked me what I wanted to do after I graduated. After hearing of my interest in science policy, she asked, "Are you sure that the day-to-day experience of science policy professionals appeals to you?" I couldn't come up with an answer that felt satisfying; all I could say was that I



**"I'd spent years
dreaming up potential jobs
based on my interests."**

thought the work would be important. "You should take on a short-term position to find out if you'll like it," she advised. Suddenly, it hit me that I'd spent years dreaming up potential jobs based on my interests, but I didn't have any experience to know whether I'd like any of them.

Over the past 11 months, I've learned that her advice to give a new career path a try was spot on. I'd gone into my company internship expecting to do standard experiments devoid of creative thought. I was prepared to get bored quickly. But there's more creativity in industry than I imagined. Few products actually make it to market, and my company is willing to try different ideas and experimentation methods, even if

they ultimately lead to a dead-end. I have been surprised by how much I've enjoyed my work.

My experience at the company has been so positive that my manager and I agreed to extend my internship to a full year. My Ph.D. advisers even allowed me to include some of the experiments I performed at the company in my dissertation.

I successfully defended my Ph.D. this month, and I am on the hunt for my first post-Ph.D. job. I've been applying for scientist positions in industry—now convinced that industry is where I would like to work after all. On the side, I hope to volunteer for groups that are working in the realm of science policy. I haven't given up on that dream, but I understand that you never know whether you'll like something until you try it out. ■

Tess Torregrosa recently completed her Ph.D. at Northeastern University in Boston, and she is in the final month of her internship at Biogen. Do you have an interesting career story to share? Send it to SciCareerEditor@aaas.org.